

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY**Type of Search****Vendors and cost where applicable**

Searcher: B. D. Schneider NA Sequence (#) 1 STN _____
 Searcher Phone #: 308-4292 AA Sequence (#) 1 Dialog _____
 Searcher Location: CMC 6.103 Structure (#) _____ Questel/Orbit _____
 Date Searcher Picked Up: _____ Bibliographic _____ Dr. Link _____
 Date Completed: 5/22 Litigation _____ Lexis/Nexis _____
 Searcher Prep & Review Time: 10 Fulltext _____ Sequence Systems Cambridge
 Clerical Prep Time: _____ Patent Family _____ WWW/Internet _____
 Online Time: 6 Other _____ Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 22:05:28 ; Search time 39 Seconds
(without alignments)
1303.977 Million cell updates/sec

Title: US-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELGVRNLISCPVWTFGT.....VRHLLSDPSGAVMVRAYLER 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1909.5	69.6	526	T07082	lycopene epsilon-c
2	749.5	27.3	500	S72505	lycopene beta-cycl
3	749.5	27.3	500	S66350	lycopene beta-cycl
4	737	26.9	500	S72506	lycopene beta-cycl
5	729	26.6	500	S66349	lycopene beta-cycl
6	705.5	25.7	498	S71511	capsanthin/capsoru
7	650.5	23.7	471	S51511	lycopene cyclase -
8	429.5	15.7	410	D75475	UbiH family protei
9	151.5	5.5	394	F81212	probable monooxyge
10	146.5	5.3	394	A81789	probable squalene
11	137.5	5.0	506	T07942	hypothetical prote
12	134.5	4.9	408	D70549	crty protein - Erw
13	125.5	4.6	382	C37802	lycopene cyclase -
14	125	4.6	386	S52981	bacteriochlorophyl
15	123	4.5	393	A75205	squalene monooxyge
16	116.5	4.2	408	H87193	probable FAD-link
17	116.5	4.2	408	T07940	NADH oxidase (H2O-
18	116.5	4.2	457	JC4541	hypothetical prote
19	115	4.2	704	AE2446	probable glutamate
20	113	4.1	476	D71200	probable lycopene
21	112.5	4.1	405	T37022	reverse gyrase (in
22	112.5	4.1	1613	S64488	probable squalene
23	112	4.1	518	T07940	2-octaprenyl-6-met
24	111.5	4.1	452	AE3596	hypothetical prote
25	111.5	4.1	456	D84202	l-aspartate oxidas
26	111.5	4.1	550	B82074	probable squalene
27	111	4.0	517	T51363	probable squalene
28	110.5	4.0	381	H69096	bacteriochlorophyl
29	110.5	4.0	382	S52585	lycopene cyclase -

30	109.5	4.0	391	2	D64366	hypothetical prote
31	108.5	4.0	291	2	H97090	thioresoxin reduct
32	107	3.9	1206	2	AG2140	hypothetical prote
33	106.5	3.9	372	2	A75185	cell division prot
34	106	3.9	431	2	H97805	ubiH protein (EC 1
35	106	3.9	531	2	T52462	hypothetical prote
36	105.5	3.8	410	2	H69307	bacteriochlorophyl
37	105.5	3.8	414	1	S74699	probable 2-octapre
38	105	3.8	393	2	C71240	hypothetical prote
39	105	3.8	435	2	F89898	glucose-inhibited
40	104.5	3.8	302	2	G71660	hypothetical prote
41	104.5	3.8	430	2	T34627	probable electron
42	104.5	3.8	437	2	H95409	probable oxidoredu
43	104.5	3.8	516	2	T51364	probable squalene
44	104	3.8	659	2	H69421	heterodisulfide re
45	103.5	3.8	481	2	H97533	dihydrolipoamide d

ALIGNMENTS

RESULT 1
T07082
lycopene epsilon-cyclase (EC 5.5.1.1) - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 19-Apr-2002
C:Accession: T07082
R:Ronen, G.; Cohen, M.; Zamir, D.; Hirschberg, J.
Submitted to the EMBL Data Library, July 1997
A:Description: Regulation of expression of the gene for lycopene epsilon cyclase dur
A:Reference number: Z15905
A:Accession: T07082
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-526 <R0N>
A:Cross-references: EMBL:Y14387; NID:el285211; PIDN:CAA74745.1; PID:el285212
A:Experimental source: cultivar VF36; leaf
C:Genetic:
A:Gene: CrtL-e-1
A:Map position: 12
C:Function:
A:Description: converts lycopene to delta-carotene
A:Superfamily: tomato lycopene beta-cyclase
C:Keywords: intramolecular lyase; intramolecular oxidoreductase; isomerase

Query Match	69.6%	Score	1909.5	DB 2	Length	526			
Best Local Similarity	68.5%	Pred. No.	8.8e-144						
Matches	368	Conservative	68	Mismatches	78	Indels	23	Gaps	6
QY	1	MELGVRNLISCPV-----WTFGRNLSSK---LAYNIHRYGSSCRVDFQVRADGG	50						
DB	1	MECVGVQNVGAMVLRPRLNWSGG--ELCQKSIFLAY--EQYESK-----NSS	48						
QY	51	SGSRSSVAYKEGFVDEEDFIRAGGSELFVOMQOTKSEKQAKLADKLPPIPFGESVMDL	110						
DB	49	SGSDSCVVDKEDFADEEDYIRAGGSQVLFVQMQKMDQSKLSDELRLQISAGOTVLDL	108						
QY	111	VVIGCGPAGLSAFAAKLGLKVGLIGDPLPFTNNYGVWEDEFKDLGLERCIEHAWKDTI	170						
DB	109	VVIGCGPAGLAAESAFLGLNVGLVGPDLPTNNYGVWEDEFKDLGLQACIEHVRDTI	168						
QY	171	VYLDNDAPVLLIGRAYGRVSRHLLHEELLKRCVSGSVILDSKVRITERAGDGHSLVVCEN	230						
DB	169	VYLDDEPILIGRAYGRVSRHFLHEELLKRCVAGVLYLNSKVDRIEATNGQSLVECEG	228						
QY	231	EIFIFCRLATVAGSAAKSLLEYEVGGPRVCVQTAYGVEVEVENPYDNLAVFMDYRDY	290						
DB	229	DVVFICRFVTVASGAASCKFLQYELGSPRVSVQTAYGVEVDNPNFDPSPFLVMDYRDY	288						
QY	291	MOOKLQCSSEETPTFLVYVMPNPTFLFFETCTCLASDKAMPFDLLKKRLMSLKLTLGIQVT	350						
DB	289	LRHDAQSLEAKYPTFLYAMPSPTRVFFETCTCLASDKAMPFDLLKKLMLRLNTLGLVRIK	348						

QY 14 PVWTFGRNLSSKLAYNIHRYGSSCRVDFQVRADGGGSRSSVAYKEGFVDEEDFIKAG 73

A;Residues: 1-500 <PEC>
A;Cross-references: EMBL:X81787; NID:gl006689; PTD:gt006690
C;Superfamily: tomato lycopene beta-cyclase
C;Keywords: intramolecular lyase; isomerase

QY 14 PVWTFGRNLSSKLAYNIHRYGSSCRVDFQVRADGGGSRSSVAYKEGFVDEEDFIKAG 73

Db 16 PVHGFVSR--ASSFNSVPRHFGSR-----KICENWNGKV-----CVKAK 53
QY 74 GSELL-FYQMOOTKSMEOAKLADKLPIPPGCVSMDLVVIGCGPAGLSLAAEAKLGLK 132
Db 54 SSALLELVETKKNLDFELPMYD-----PSKGLVVDLAVVGGGPAGLAVAAQVSEAGLS 108
QY 133 VGLI--GPDLPFTNNYGVWEDEFKDLGLERCIEHAKDTIYVLDNDAPVLIGRAYGRYSR 190
Db 109 VVSDPSFKLIWPNNGYVWVDEFEAMDLDCLDATWSGTVVYIDNTTKDLDPRYGRVNR 168
QY 191 HLLHEELLKRVESGVSYLDSKVRITERBAGDGHSLVWCNEIFIPCRLATVASGAASGLK 250
Db 169 KQLASKMOKCILNGVFXHAKVIVKVIHE-EAKSMLICNDGVTIQATVVLDTATG-FSRCL 226
QY 251 LEYEVGGP-RVCVOTAYGVEVENNPVDPNLMPMDYRD-YMOOKLOCSE--EYPTFL 306
Db 227 VOYD--FYPKPGYQVAGIILAEVEHFDTSKVMIMDWRDHLGNMMLKERNKRVPTFL 284
QY 307 YVPMSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIQVTKVYEEESYIPVGGSL 366
Db 285 YAMPFSSNKIFLETSVLARPGLRMDDDIQRVAVRLNHLGKVKISIEDEHCVIPMGSL 344
QY 367 PNTQKKNLACAAASVHPATGYSVRSLSAPKYSVIAKILKQDNSAYVVSQSSAVN 426
Db 345 PVIPQVVGTTGAGLVHPSTGYMVARTLAAAPVAVNAIHYLGSEK--LLGNLSAA- 401
QY 427 ISMQAWSLWPKERKORAFPLFGLLEIVOLDIEATRTFFRFLPTWMMWGLGSL 486
Db 402 ----VWKDLWPIERRRORREFFCGMDILLKLDLPATRRFFDAFFDLPRYWHGFLSR 457
QY 487 SFDLVFSMYMFLVAPNSMR-----MSLVRHLLSD 516
Db 458 LPELIFFGLSLFRASNTSRIEIMTKGTLPVNMNLLQD 498

RESULT 5
S66349
lycopene beta-cyclase (EC 5.5.1.1) - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 28-Jul-2000
C:Accession: S66349
R:Pecker, I.; Gabbay, R.; Cunningham Jr., F. X.; Hirschberg, J.
Plant Mol. Biol. 30, 807-819, 1996
A:Title: Cloning and characterization of the cDNA for lycopene beta-cyclase from tomato
A:Reference number: S66349; MUID:96194462; PMID:8624411
A:Accession: S66349
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-500 <PEC>
A:Cross-references: EMBL:X81787
C:Superfamily: tomato lycopene beta-cyclase
C:Keywords: intramolecular lyase; isomerase

Query Match 26.6%; Score 729; DB 2; Length 500;
Best Local Similarity 35.1%; Pred. No. 6.3e-50;
Matches 163; Conservative 103; Mismatches 165; Indels 34; Gaps 12;

QY 70 IRAGSELL-FVQMOTKSMEOAKLADKLPIPPGCVSMDLVVIGCGPAGLSLAAEAK 128
Db 50 VRAKSSALLELVETKKNLDFELPMYD-----PSKGLVVDLAVVGGGPAGLAVAAQVSE 104
QY 129 LGLKVLGI--GPDLPFTNNYGVWEDEFKDLGLERCIEHAKDTIYVLDNDAPVLIGRAY 186
Db 105 AGLSVYSIDPSKLIWPNNGYVWVDEFEAMDLDCLDATWSGVVYIDNTTKDLDPRYG 164
QY 187 RYVSRHLLHEELLKRCVSGVSYLDSKVRITERBAGDGHSLVWCNEIFIPCRLATVASGA 246
Db 165 RVNRKOLSKMOKCILNGVFXHAKVIVKVIHE-EAKSMLICNDGVTIQATVVLDTATG-F 222
QY 247 SGKLLLEYEVGGP-RVCVOTAYGVEVENNPVDPNLMPMDYRD-YMOOKLOCSE--EY 302
Db 223 SRLCLVQYD--KPYKPGYQVAGIILAEVEHFDTSKVMIMDWRDHLGNMMLKERNKRV 280

QY 303 PTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIQVTKVYEEESYIPV 362
Db 281 PTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIQVTKVYEEESYIPV 340
QY 363 GGSLLPNTQKKNLAFCAASVHPATGYSVRSLSAPKYSVIAKILKQDNSAYVVSQSS 422
Db 341 GGSLLPNTQKKNLAFCAASVHPATGYSVRSLSAPKYSVIAKILKQDNSAYVVSQSS 398
QY 423 SAVNISMQAWSLWPKERKORAFPLFGLLEIVOLDIEATRTFFRFLPTWMMWGLG 482
Db 399 SAA-----VWKDLWPIERRRORREFFCGMDILLKLDLPATRRFFDAFFDLPRYWHGFLS 453
QY 483' SLSLSDFDLVFSMYMFLVAPNSMR-----MSLVRHLLSD 516
Db 454 SRLYLPELIFFGLSLFRASNTSRIEIMTKGTLPVNMNLLQD 498

RESULT 6

S71511
capsanthin/capsorubin synthase (EC 5.5.1.1) - pepper
N:Alternate names: chromoplast-specific protein
N:Contains: capsorubin synthase; capsanthin synthase
C:Species: Capsicum annuum (pepper)
C:Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 28-Jul-2000
C:Accession: S71511; JC2141
R:Bouvier, F.; Hugueney, P.; d'Harlingue, A.; Kuntz, M.; Camara, B.
Plant J. 6, 45-54, 1994
A:Title: Xanthophyll biosynthesis in chromoplasts: Isolation and molecular cloning of
A:Reference number: S71511; MUID:95004653; PMID:7920703
A:Accession: S71511
A:Molecule type: mRNA
A:Residues: 1-498 <BOU>
A:Cross-references: EMBL:X76165; NID:9522119; PIDN:CAA53759.1; PID:9522120
A:Experimental source: cv. Lamuyo; developmental stage ripening fruits
R:Deruere, J.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.
Biochem. Biophys. Res. Commun. 199, 1144-1150, 1994
A:Title: Structure and expression of two plant genes encoding chromoplast-specific
A:Reference number: JC2140; MUID:94197697; PMID:8147854
A:Accession: JC2141
A:Molecule type: DNA
A:Residues: 1-498 <DER>
A:Cross-references: EMBL:X77289; NID:9468747; PIDN:CAA54495.1; PID:9468748
R:Deruere, J.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.
Biochem. Biophys. Res. Commun. 201, 486, 1994
A:Reference number: A55775
A:Contents: annotation; erratum
A:Note: the legends of the nucleotide sequence figures were reversed in publication

A:Gene: CCS
C:Complex: monomer
C:Function:
A:Description: catalyses the conversion of the ubiquitous 5,6-epoxycartenoids, anthra
A:Pathway: carotenoid biosynthesis
C:Superfamily: tomato lycopene beta-cyclase
C:Keywords: chromoplast; intramolecular lyase; isomerase

Query Match 25.7%; Score 705.5; DB 2; Length 498;
Best Local Similarity 38.6%; Pred. No. 4.7e-48;
Matches 158; Conservative 79; Mismatches 153; Indels 19; Gaps 9;

QY 109 DLVYVIGCGPAGLSLAAEAKLGLKVLIGDLPFT---NNGVWEDEFKDLGLERCIEHA 165
Db 83 DVIITIGPAGLRLAEQVSKYGIKVCVDPFS-PLSMFNNYGVWVDEFEKLGLEDCLDHK 141
QY 166 WKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVRITERBAGDGHSL 225
Db 142 WPVSCVHISHDKYKYLDRPGRVSRKLLKLLNSCVENRVKFKAKVLRKHE-EPESS 200
QY 226 VVCNEIFIPCRLATVASGAASGLLEYEVGGP-RVCVOTAYGVEVENNPVDPNLMPV 284
Db 201 IVCDGGRKISGLTVDASGVAS-DFEYD--KPRNHGVQVAGHILAEVDNHPDLDKML 257

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <TET>
A;Cross-references: GB:AE002389; GB:AE002098; NID:g7225537; PIDN:AAF40768.1; PID:g722

us-09-701-395a-23.rpr

Thu May 22 09:39:18 2003

A: Experimental source: serogroup B, strain MC58

C: Genetics:

A: Gene: NMB0323

Query Match 5.5%; Score 151.5; DB 2; Length 394;
 Best Local Similarity 21.7%; Pred. No. 0.00036;
 Matches 84; Conservative 57; Mismatches 133; Indels 113; Gaps 17;

QY 109 DLVVIGCGPAGLSAAEAKGLKVLII-----GPDLPFTN-----NYG 147
 Db 6 DILVVGAGPAGLSFAAEAGLGLKVLIIERSPLTVLQPPVDGRIALTFRSREIMORLG 65
 QY 148 VW-----EDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGR-----AYGRVS-- 189
 Db 66 MWDKIPENE-----IYPLRDAKVLNGRSYQLHFPOPTGEARPADC 107
 QY 190 -RHLHEELLKRCVSGSYLDS-----KVERITEAGDGHSLVCENEIFIPCLATVA 242
 Db 108 LGLYLSNHNIRAAEYVVSQDNVSLITDIVVKEVKTSDNEAQVILENGKILTARLLAA 167
 QY 243 SGAASGKLEXYEVGPRVCQTAYGVEVEVNNPYDNLVMDYRDYMOOKLQCSSEY 302
 Db 168 DSRFS-----QTRRQLGIGSSDMHDYSTRMFVCR-----MKHTLSNQHTAY 207
 QY 303 PTFLY--VMPMSPTRLFFETCL-----ASKDAMPFDL-----LKRKLSRLKTLGIQV 349
 Db 208 ECFHYGRTIALPLEEHLNTVITVDTKINSVQNLSPLELAASVKQLKGRIGDMEL-V 266
 QY 350 TKVYEESWYIPVGGSLPNT--EOKNLAFGAASVHPATGYSVVRSISEAPKYASVI-- 405
 Db 267 SSIH-----HYPLVGMIAKRFYGRSALIGDAAGVHPVTAHGFNLGLSSADILAKLLE 321
 QY 408 ILKQDNASVVGSSAVNISQWSS 434
 Db 318 LITAEQ-----RGQDIGASSLLEKYSS 340

RESULT 10
 A: Experimental source: serogroup A, strain Z2491
 C: Genetics: NMB2164
 A: Gene: NMB2164

Query Match 5.3%; Score 146.5; DB 2; Length 394;
 Best Local Similarity 21.5%; Pred. No. 0.00095;
 Matches 83; Conservative 56; Mismatches 130; Indels 117; Gaps 16;

QY 109 DLVVIGCGPAGLSAAEAKGLKVLII-----GPDLPFTN-----NYG 147
 Db 6 DILVVGAGPAGLSFAAEAGLGLKVLIIERSPLTVLQPPVDGRIALTFRSREIMORLG 65
 QY 148 VW-----EDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGR-----AYGRVS-- 189
 Db 66 MWDKIPENE-----IYPLRDAKVLNGRSYQLHFPOPTGEARPADC 107
 QY 190 -RHLHEELLKRCVSGSYLDS-----KVERITEAGDGHSLVCENEIFIPCLATVA 242
 Db 108 LGLYLSNHNIRAAEYVVSQDNVSLITDIVVKEVKTSDNEAQVILENGKILTARLLAA 167

QY 243 SGAASGKLEXYEVGPRVCQTAYGVEVEVNNPYDNLVMDYRDYMOOKLQCSSEY 302
 Db 168 DSRFS-----QTRRQLGIGSSDMHDYSTRMFVCR-----MKHTLSNQHTAY 207
 QY 303 PTFLY--VMPMSPTRLFFETCL-----ASKDAMPFDL-----LKRKLSRLKTLGIQV 349
 Db 208 ECFHYGRTIALPLEEHLNTVITVDTKINSVQNLSPLELAASVKQLKGRIGDMEL-V 266
 QY 350 TKVYEESWYIPVGGSLPNT--EOKNLAFGAASVHPATGYSVVRSISEAPKYASVI-- 405
 Db 267 SSIH-----HYPLVGMIAKRFYGRSALIGDAAGVHPVTAHGFNLGLSSADILAKLLE 321
 QY 406 -----AKILKQDNASVVGSGQ 421
 Db 322 AEQGDIGAGSALLEKYSKHLHAQ 347

RESULT 11

T07942

probable squalene monooxygenase (EC 1.14.99.7) Squal - rape

N: Alternate names: probable squalene epoxidase

C: Species: Brassica napus (rape)

C: Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 21-Jul-2000

C: Accession: T07942

R: Schaefer, U.A.; Reed, D.W.; Hunter, D.G.; Yao, K.; Weninger, A.M.; Tsang, E.W.T., F

Plant Mol. Biol. 39, 721-728, 1999

A: Title: An example of intron functional sliding in the gene families encoding squal

A: Reference number: Z16234; MUID: 99277589; PMID: 10350086

A: Accession: T07942

A: Status: translated from GB/EMBL/DBJ

A: Molecule type: mRNA

A: Residues: 1-506 <SCH>

A: Cross-references: EMBL:AJ005931; NID:G3123353; PIDN:CAA06773.1; PID:G3123354

A: Experimental source: cv. Westar; developmental stage: greening-etiolated; tissue ty

C: Genetics:

A: Gene: Sqp1

A: Introns: 480/3

C: Keywords: FAD; oxidoreductase

Query Match 5.0%; Score 137.5; DB 2; Length 506;

Best Local Similarity 24.3%; Pred. No. 0.0071;

Matches 108; Conservative 52; Mismatches 145; Indels 139; Gaps 25;

QY 89 EKOAKLADKLPPIFGESVMDLVIGCGPAGLSAAEAKGLKVLIGLIPDL--PFTNNY 146

Db 30 KKVAKLPDAATEVR-RDGDADTVIAGVGGSAALAKDGRVHVTERDMREPRV-- 85

QY 147 GWDEDF-----KDLGLERCIE-----HAKDRIYV-----DNDAPV-LI 181

Db 86 --MMGEFMQPGGRLLLSKLGLEDCELDQIATGLAVYKDGKALVSPFENDFPYPT 143

QY 182 GRAY--GRVSRHL-----LHEELKRCVE-----SGVSYLDSKVERITEAGDGH 223

QY 144 GRAFYNGRFVORLROKASSLFTVQLEGGTVKSLIEKGVIKVYKNS-----AGE-- 194

Db 224 SLVVCENEIFIPCLATVAGSAAS-----GKILLEYGGRVCQTAYGVEVEVE 273

QY 195 -----ETTAFAP--ITVVCDCGYSNLRSSVNDNNAEYISYOVG-----YVSKNCQLE-- 239

Db 274 NNPDNLVMDYRDYMOOKLQCSSEYPTFLVYMPMSPTRLFFETCLASQDAM---- 329

QY 240 -----DPEKLKIMSKPSFTMLYQISSTDVRCVMEIFPGNIPISNGEMAVYKNTMAPQV 295

Db 330 PFDLLKRLKLSRLKTLGIQV---TKVYEESWYIPVGGSLPNTPEOKNLAFGAASVHP 385

QY 296 PPEL--RKIFLKGIDEGAQIKAMPTKRMEATLS-----EKQGVIVLGDADFNRHP 343

Db 386 --ATGYSVW-----RSLSEAPKYASVIKILKQDNASVYV--SQSSAVNIS 428

QY 344 ATASGMVVLSDILIRLLQLPLNLSNDANKVSEVI-----KSFYVIRKPMSTATNLT 396

Db 429 MQAWSSLW-----PKERKQRAF 446

Db 397 GNAFSQVLIATDEAKEAMROGCF 420

RESULT 12

D70549

Hypothetical protein Rv0561c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
R:Accession: D70549

R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Rajandream, M.A.; Rogers, K.; Devlin, K.; Feltri, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nature 393, 537-544, 1998
A:Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A:Title: Deciphering the Biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-408 <CO>
A:Cross-references: GB:295558; GB:AL123456; NID:93261781; PIDN:CAB08972.1; PID:g2114023
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0561c
C:Superfamily: fixC protein

Query Match 4.9%; Score 134.5; DB 2; Length 408;
Best Local Similarity 23.8%; Pred. No. 0.009;
Matches 91; Conservative 45; Mismatches 130; Indels 117; Gaps 20;

QY 109 DLVVGCGPAGLSAAEAKGLKVLGIGP-----DLPTN-----NYGVW- 149

Db 8 DVVVVGAGPAGSAAWAAARAGRDVLVDATFPRDKPCGDLGPRAVAELHQLGLKWL 67

QY 150 EDEFKDLGLERC-----IEHAWKDTIVLDNDAPVLIGRAYG-RVSRHLLHEELKRCVE 203

Db 68 ADHIRHRLGRLMSFGGEVEVDW-----PGSPFSGYSAVARLELDDRIKVAED 116

QY 204 SGVSYL-----DS--KVERITEAGDCHSLVVCENEIFIPCLATVASCAS--GK 249

Db 117 TGARMLLAKAVAVHDSRRVSLTLA-DGTE-----VGCQLIVADGARSPLGR 166

QY 250 LLEVEVGPRVCVOTAYGVEVE-VENNPY--DENLVMFYMDYRDYMOOKLQCCSEEEYPTFL 306

Db 167 KL-----GRWRHRETVYGVAVGYLSTAYSDDPWLTSHLELRSPDCAVL-----PGYG 214

QY 307 YVMPSPTRLFFETCTCLAKDAMPDLKRLKMSRLKTLGIGVTKVVEEWSY----- 359

Db 215 WIFPLNGEYVIGVGL-STSRPADLALRPLIS-----YYTDLRDEMGFTGQPRV 266

QY 360 -----IPVGSPLNTEQKN-LAFGAASVMVHPATGYSVVRSLSSEAPKYASVIKILKODNS 414

Db 267 SSALLPMGAVSGVAGSNMMLIGDAAACVPLNGEIDYGLE----- 308

QY 415 AYVVGQSSAVNISMQWSSSLWP 437

Db 309 -----TGLRAAELDSRLARLWP 327

RESULT 13

C37802

crty protein - Erwinia uredovora
C:Species: Erwinia uredovora
C>Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 20-Jun-2000
R:Accession: C37802

R:Misawa, N.; Nakagawa, M.; Kobayashi, K.; Yamano, S.; Izawa, Y.; Nakamura, K.; Harashim J. Bacteriol. 172, 6704-6712, 1990
A:Title: Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional genomics
A:Reference number: A37802; PMID:91072214; PMID:2254247

A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-382 <MS>

A:Cross-references: GB:D90087; NID:g216681; PIDN:BAA14126.1; PID:g216684
C:Superfamily: Erwinia dycopene cyclase

Query Match

Best Local Similarity 4.6%; Score 125.5; DB 2; Length 386;
Matches 79; Conservative 74; Mismatches 185; Indels 57; Gaps 15;

QY 109 DLVVGCGPAG--LSLAAEAAKGLKVLGIGPDLPTNNYGVWEDEFKDL-----GLER 160

Db 6 DLILVGGGLANGLIARLQRYPOLNLLIEAGQPGNGNHTWSFHEDDLPQGHAWLAPL 62

QY 161 CIEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEELKRCVSGVSYLDSKVRIEYTA 219

Db 65 LVVHWPDYQVRPTRRRKNSGYCITTSQRF--AEVLQR--QFGPHLMVATAVEV--- 117

QY 220 GDGHSLLVVCENEIIPCLATVASCASGKLLVEYVGPRVCVOTAYGVEVEVENNPYDP 279

Db 118 -NAESVRLKKGQV-IGARAVIDGGRVYANALS-----VGFOAFICQENRL-SHPHGL 167

QY 280 NLVMFYMDYRDYMOOKLQCCSEEEYPTFLVMPSPTRLFFETCTCLAKDAMPDLKRLKLM 339

Db 168 SSPIMD-----ATVDQNGYRFVYSLPSLPTLLIEDTHYIDNATLDPECARONIC 219

QY 340 SRLKTLGIGVTKVVEEWSYIPV---GGSPLNTEQKNLA-FGAASVMVHPATGYSVVRS 395

Db 220 DYAAOQWQLOTLREEOGALPITLGNADAFWQORPACSGRLAGLHPTTGYSLPLAV 279

QY 396 SEAPKYASVIKILKQDNSAVVSGSSAVNISMQWSSSLWPKRKRQRAFF-LFGLLELI 454

Db 280 AVADR-----LSALDVTTSASIIHAITHFAERWQOQGGFFMLNMLF 322

QY 455 VOLDIEATRTFFRFFRLPTMMWGLGSSLSFD 489

Db 323 LAGPADSRVVMORFVGLPDLIAREYAGKLITLD 357

RESULT 14

S52981

lycopene cyclase - Erwinia herbicola
C:Species: Erwinia herbicola
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 10-Dec-1999

R:Accession: S52981
R:Rundle, B.; Alberti, M.; Nieveinstein, V.; Beyer, P.; Kleinig, H.; Armstrong, G.A.; Mol. Gen. Genet. 245, 406-416, 1994
A:Title: Functional assignment of Erwinia herbicola Eho 10 carotenoid genes expressed in Escherichia coli
A:Reference number: S52976; PMID:95107236; PMID:7808389

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-386 <HUN>

A:Cross-references: EMBL:M07280; NID:g148404; PIDN:AAA64980.1; PID:g148411
A:Start codon: GTC
C:Superfamily: Erwinia dycopene cyclase

Query Match

Best Local Similarity 4.6%; Score 125; DB 2; Length 386;
Matches 86; Conservative 62; Mismatches 163; Indels 112; Gaps 18;

QY 109 DLVVGCGPAG--GL---SLAAEAAKGLKVLGIGPDLPTNNYGVWEDEF---KDLGLERC 161

Db 3 DLILVGGGLANGLIARLQRYPOLNLLIEAGQPGNGNHTWSFHEDDLPQGHAWLAPL 62

QY 162 IEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEELKRCVSGVSYLDSKVRIEYTA 219

Db 63 VAHAWPGVEVQFPD-----LRRRLARG--YYSITSERFAEALH 98

QY 220 ---GDGHSLLVVCENEIIPCLATV-----ASGAA--SGKLEVEGPRVCVOTAYG 267

Db 99 QALGE-----NIWLCVSSEVLPSNVRVLANGALLAGAVIDGRGVYASSAMOTGYQ 149

QY 268 VEVEVE---NNPYDPNLMVMDYRDYMOOKLQCCSEEEYPTFLVMPSPTRLFFETCTLA 324

Job time : 44 secs

Db 150 LFLGQOMRLTQPHGLVVPILMDATVAQQQYR-----FVYTLPLSADTLLIEDTRYA 201
QY 325 SKDAMPFDLLKRLKMSRLKTLGIQVTKVVEEWSYIPV--GGSL-----PNTQKNLA 375
Db 202 NYPQRDDNALRQVTDYAHSGWQIAQLEREETGCLPITLAGDIOALWADAPCVPRS--- 258
QY 376 FGAASWVHPATGYS-----VYRSLSSEAPKYASVIKILKODNSAYVVGSSAYVNISM 429
Db 259 -GMRAGLFHPTTGYSLPLAVALADATADSPRLGSV-----PLYQLTR 299
QY 430 QAWSSLPKE---RKQRAFFLEGLLELIVQLDIEATRTFFRFLPTWMMWGLGSSLS 486
Db 300 QFAERHWRQGFRLNRLMLAGRE-----ENRWYMQRFYGLPEPTVERFYAGRLS 352
QY 487 SFD 489
Db 353 LFD 355

RESULT 15
A75205
bacteriochlorophyll synthase 43 kda chain or geranylgeranyl hydrogenase PAB0109 - Pyrococcus
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: A75205
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: A75205
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <KAW>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49088.1; PID:g545759
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: bchP; PAB0109
C:Superfamily: hypothetical protein MJ0532

Query Match 4.5%; Score 123; DB 2; Length 393;
Best Local Similarity 22.3%; Pred No. 0.07;
Matches 86; Conservative 55; Mismatches 181; Indels 64; Gaps 15;

QY 109 DLVWIGCGPAGLSAAEAKLGLKVLIGP-----DLPFTNNYGVWEDEFKDLGL-----ERC 161
Db 4 DVVWVSGVAGPTVARDVAKAGFSVLLVDKAAIGTPKQCAEGITIEVFKEFDIPYDKRF 63
QY 162 IEHAWKDTIVLDNAPVLIGRAYGRVSRHLHEELLKRCVSGVSYLDSKVERITEAGD 221
Db 64 INREIYAKIYSPSGYTAEL--RYKHVSGVILERKVFDRKMLAYYAAKAGAEVLARTEVVD 121
QY 222 GHSLVVCEN-----EIFPCRLATVAGSAAS--GKLLVEYVGGPRVCVQTAY 266
Db 122 ----VIRNGKIVGRVRAKEGEPLEIEAKVIVAADGVESTIARKAGINTYAPHEFDSGY 177
QY 267 GVEVEVNNPDYDNLVMDYDMQKLOCSEEEYPT-FLXVMPMSPTLFLFEETCLAS 325
Db 178 EYEMLIIEG--YDPDLI-----HLFFGNEVAPRGVWIFPKDEDR---ANVGIGI 221
QY 326 KDAMPFDLLKRLKMSRLKTLGIQVTKVVEEWSYIPVGGSLPNTQKN-LAFGAASWVH 384
Db 222 ASDHP-QTAKYYLDKWLKNNIPMRKILEVNVGLVPVGGFVKRELKVDNVLVVGDAARQYN 280
QY 385 PATGYSVVRSLSEAPKYASVIKILKODNSAYVVGSSAYVNISMQAWSSLP-----K 438
Db 281 PVHGGMKEMKAASIAAKWIKALEENLELL-----KNYSEEWRTGEPMEKLLK 333
QY 439 ERKQRAFFLGLLEIVQL-----DIE 460
Db 334 LRKAMEKLDELDVFLVQVLSGTDLE 359

Search completed: May 21, 2003, 22:12:01

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:29:03 ; Search time 233 Seconds
(without alignments)
2997.959 Million cell updates/sec

Title: US-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRNLISCPWTFGT.....VRHLLSPSGAVMYRAYLER 529

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 828747 seqs, 660231138 residues
Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09701395/runat_15052003_165754_3440/app_query.fasta.1.711
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LISP=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09701395 -SCGN_1_135 -runat_15052003_165754_3440
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2743	100.0	1898	10	US-09-323-998D-22
2	2728	99.5	1848	10	US-09-323-998D-46
3	1959.5	71.4	1860	10	US-09-323-998D-1
4	1478.5	53.9	1378	10	US-09-323-998D-24

5	1394	50.8	1272	12	US-10-041-472-3	Sequence 3, Appli
6	1202	43.8	960	12	US-10-041-472-2	Sequence 2, Appli
7	739	26.9	1590	12	US-10-041-472-4	Sequence 4, Appli
8	173	6.3	1185	9	US-10-166-037-2	Sequence 2, Appli
9	129.5	4.7	1149	9	US-09-941-947A-29	Sequence 29, Appli
10	123	4.5	1593	9	US-09-938-842A-1578	Sequence 1578, Ap
11	115.5	4.2	1272	9	US-09-738-626-529	Sequence 529, App
12	113	4.1	1149	10	US-09-547-267-8	Sequence 8, Appli
13	113	4.1	8625	9	US-09-920-923-1	Sequence 1, Appli
14	113	4.1	11233	9	US-09-920-923-27	Sequence 27, Appli
15	111.5	4.1	536165	9	US-09-939-964-1	Sequence 1, Appli
16	111	4.0	1778	10	US-09-887-576-634	Sequence 634, App
17	106	3.9	6796	7	US-08-781-986A-2	Sequence 2, Appli
18	104.5	3.8	1551	9	US-09-938-842A-1094	Sequence 1094, Ap
19	103	3.8	2326	9	US-10-025-380-1103	Sequence 1103, Ap
20	103	3.8	2326	10	US-09-922-217-1103	Sequence 1103, Ap
21	102	3.7	2025	10	US-09-070-927A-439	Sequence 439, App
22	101	3.7	1308	10	US-09-815-242-6599	Sequence 6599, Ap
23	101	3.7	1999	10	US-09-917-800A-1706	Sequence 1706, Ap
24	99.5	3.6	1704	10	US-09-815-242-9465	Sequence 9465, Ap
25	99	3.6	1407	9	US-09-738-626-764	Sequence 764, App
26	99	3.6	2389	9	US-10-010-920-1	Sequence 1, Appli
27	99	3.6	2389	9	US-10-270-877-1	Sequence 1, Appli
28	99	3.6	2389	9	US-10-270-877-1	Sequence 1, Appli
29	99	3.6	2389	9	US-10-008-721-1	Sequence 1, Appli
30	98	3.6	549	9	US-09-991-936-256	Sequence 256, App
31	98	3.6	1179	10	US-09-815-242-7698	Sequence 7698, Ap
32	98	3.6	1303	10	US-09-815-242-4838	Sequence 4838, Ap
33	98	3.6	1308	10	US-09-815-242-8761	Sequence 8761, Ap
34	97.5	3.6	1551	9	US-09-938-842A-2698	Sequence 2698, Ap
35	97.5	3.6	3014	9	US-09-933-261-1	Sequence 1, Appli
36	97.5	3.6	3014	9	US-10-256-702-1	Sequence 1, Appli
37	97.5	3.6	3309400	9	US-09-738-626-1	Sequence 1, Appli
38	97	3.5	1385	9	US-10-005-530-40	Sequence 40, Appli
39	95.5	3.5	2196	10	US-09-917-800A-1469	Sequence 1469, Ap
40	95	3.5	1260	10	US-09-815-242-9700	Sequence 9700, Ap
41	95	3.5	2206	9	US-10-108-605-258	Sequence 258, App
42	95	3.5	3989	10	US-09-070-927A-359	Sequence 359, App
43	94.5	3.4	2493	9	US-09-938-842A-322	Sequence 322, App
44	93.5	3.4	2169	9	US-09-906-419-8	Sequence 8, Appli
45	93.5	3.4	2186	12	US-10-062-254-91	Sequence 91, Appli

ALIGNMENTS

RESULT 1
US-09-323-998D-22
; Sequence 22, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Adonis palaestina
US-09-323-998D-22

Alignment Scores:

Pred. No.:	0	Length:	1898
Score:	2743.00	Matches:	529
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-701-395A-23 (1-529) x US-09-323-998D-22 (1-1898)

QY	1	MetGluLeuLeuGlyValArgAsnLeuIleSerCysProValTrpThrPheGlyThr	20
Db	113	ATGGAACTACTTGGTGTTCCGAACCTCATCTCTTCTGGCCCTGCTGGACATTTGGACAA	172
QY	21	ArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArg	40
Db	173	AGAAACCTTAGTAGTTTCAAACTAGCTTATAACATACATGATATGGTTCTTCTTGTAG	232
QY	41	ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerSerValAlaTyrLys	60
Db	233	GTAGATTTTCAAGTCAGAGCTGATGGTGGAGCGGAGTAGAAGTTCTGTGTGCTATAAA	292
QY	61	GluGlyPheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheVal	80
Db	293	GAGGGTTTGTGGATGAAGAGATTATCAAAAGCTGGTGGTCTGAGCTTTGTGTGTC	352
QY	81	GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuPro	100
Db	353	CAAAATGCAGCAACAAAGTCATGGAGAACAAGCCCAAGCTCCCGATAAAGTGCACCA	412
QY	101	IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu	120
Db	413	ATACCTTTTGGAGAATCCGTGATGGACTTGGTTGTAATAGGTGTGGACCTGCTGGTCT	472
QY	121	SerLeuAlaAlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu	140
Db	473	TCACTGGCTGCAGAAAGCTGCTAAAGCTAGGGTTGAAAGTTGGCTTATTGGTCTGATCT	532
QY	141	ProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuArg	160
Db	533	CCTTTACAAATAATTATGTGTGTGGGAGACGAGTTCAAAGATCTTGGACTTGAACGT	592
QY	161	CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu	180
Db	593	TGTAATCAGCATGCTTGGGAAGSACACATCGTATATCTTGATAATGATGCTCCTGCTCT	652
QY	181	IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLysArg	200
Db	653	ATTGGTCGTCATATGCACGAGTTATGCACATTTGCTACATGAGGAGTTGCTGAAAAG	712
QY	201	CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly	220
Db	713	TGTTGGAGTCAGTGTATCATATCTTGATTTCTAAAGTGGAAAGGATCCTGAAGCTGGT	772
QY	221	AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr	240
Db	773	GATGGCCATAGCCCTGTAGTTTGTGAAATGAGATCTTATCCCTTGCAGGCTGCTACT	832
QY	241	ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgVal	260
Db	833	GTTCATCTGGACAGCTTCAGGGAACCTTTTGGAGTATGAAGTAGGTGGCCCTCGTGTT	892
QY	261	CysValGlnThrAlaTyrGlyValGluValGluAluAsnAsnProTyrAspProAsn	280
Db	893	TGTGTCCAAACCGCTTATGGGTGGAGCTTGAGGTGGGAGCAACATCCATAGATCCCAAC	952
QY	281	LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu	300
Db	953	TTAATGGTATTATGGACTACAGAGACTATATGCAACAGAAAATTACAGTCTCGGAAGAA	1012
QY	301	GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu	320
Db	1013	GAATATCAACAATTTCTATGTATGCGCCATGTCGCCAACAGACTTTTTTTTGGAGAA	1072

QY	321	ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer	340
Db	1073	ACCTGTTGGCGCTCAAAAGATGCCATGCCATTGCATCTACTGAAGAGAAACTGATGTC	1132
QY	341	ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrrGluGluGluTrpSerTyrIle	360
Db	1133	CGATTGAAGACTCTGGGTATCCRANGTTTACAAAAGTTTATGAAGAGGAATGGTGCATATTT	1192
QY	361	ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAla	380
Db	1193	CCTGTTGGTGGTCTTTTACCAAAACACAGACAAAGAACCTAGCATTTGGTGCTGCAGCA	1252
QY	381	SerMetValHisProAlaThrGlyTyrSerValValAlaGlySerLeuSerGluAlaProLys	400
Db	1253	AGCATGTGTCATCCAGCAACAGCGCTATTCTGGTTTACGGTCTACTGTCAGAAGCTTCAAAA	1312
QY	401	TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly	420
Db	1313	TATGCTTCTGTAAATGCAAGATTTTGAAGCAAGATAAATCTGCGTATGTGGTTCCTGGA	1372
QY	421	GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg	440
Db	1373	CAAAAGTAGTCAGTAAACATTTCAATGTCAAGCATGGACGAGCTTTGGCCAAAGACGCGA	1432
QY	441	LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu	460
Db	1433	AAAGCTCAAGACAGCATCTTCTTTTGGATTAGAGCTTATTGTGCAGCTAGATATTGAA	1492
QY	461	AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTyrTrpGlyPhe	480
Db	1493	GCAACACAGAACATCTTTAGAACCCTTCTCCGCTTGCCACTTGGATGTGGTGGGGTTTC	1552
QY	481	LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu	500
Db	1553	CTTGGGTCTCACTATCATCTTTCATCTCGTCTGTGTTTTCCATGTACATGTTGTGTTTG	1612
QY	501	AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla	520
Db	1613	GCGCCAAACAGCATGAGGATGTCATCTGTGACATTTGCTTTTCAGATCCTCTCTGGTGCA	1672
QY	521	ValMetValArgAlaTyrIleGluArg	529
Db	1673	GTATTGTTAAGACTTACTCGAAGG	1699

RESULT 2

```

US-09-323-998D-46
; Sequence 46, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZATREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 1848
; TYPE: DNA
; ORGANISM: Adonis palaestina
US-09-323-998D-46

```

US-09-323-998D-46

us-09-701-395a-23.rnpb

Thu May 22 09:39:17 2003

Alignment Scores:
 Pred. No.: 0 Length: 1848
 Score: 2728.00 Matches: 524
 Percent Similarity: 100.00% Conservative: 5
 Best Local Similarity: 99.05% Mismatches: 0
 Query Match: 99.45% Indels: 0
 DB: 10 Gaps: 0

US-09-701-395a-23 (1-529) x US-09-323-998D-46 (1-1848)

QY 1 MetGluLeuGlyValArgAsnLeuLeuSerCysProValTrpThrPheGlyThr 20
 DB 116 ATGAACCTACTTGGTGTCCACCACTCACTCTTCTGCGCTCTGCGACTTTTGAACA 175
 QY 21 ArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerCysArg 40
 DB 176 AGAACCTTAGTCTCAAACTAGCTTATAACATACATCGATATGGTCTCTCTGTAGA 235
 QY 41 ValaspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLys 60
 DB 236 GTAGATTTCAGTGGAGGCTGATGGTGAAGCGGAGTAGAATCTCTGTGCTTATAAA 295
 QY 61 GluGlyPheValAspGluGluaspPheIleLysAlaGlyGlySerGluLeuLeuPheVal 80
 DB 296 GAGGTTTGTGACGAGGAGGATTTATCAAGCTGTGTGCTCTGAGCTTTGTGTC 355
 QY 81 GlnMetGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
 DB 356 CNAATGAGCAACAAGCTATGAGAAACAGCGCTGCGGATTAAGTTCACCA 415
 QY 101 IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
 DB 416 ATACCTTTCGAGAACTCTGTGAGCTTGGTGTGTAAGTGTGGACCTGTGCTCT 475
 QY 121 SerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
 DB 476 TCATGCTGCGAAGCTCTAGCTAGCTTCAAAAGTTGGCTTATGCTGCTGATCT 535
 QY 141 ProPheThrAsnAsnTyrGlyValTrpGluaspGluPheLysAspLeuGlyLeuArg 160
 DB 536 CTTTTTACAAATAATATGTTGTGTGGAGAGAGCTTCAAGATCTGAGCTTGAACGT 595
 QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180
 DB 596 TGTATCGAGATGCTTGAAGGACCATCGTATATCTGACATGATGATGCTGCTGCT 655
 QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLysArg 200
 DB 656 ATGGTCTGCTATGAGAGCTTACCGGCACTTTCGCTGATGAGAGTTCGTAAGAG 715
 QY 201 CysValGluSerGlyValSerTyrLeuaspSerLysValGluArgIleThrGluAlaGly 220
 DB 716 TGTGTCGAGTCAGCTGTATCATATCTGATTTCTAAAGTGAAGAGATCAGTGAAGCTGGT 775
 QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
 DB 776 GATGCCATAGCTCTGTAGTTTGTGAACAGCACTTTATCCCTGCGAGCTTGTCTACT 835
 QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyClyProArgVal 260
 DB 836 GTTGATCTGGAGCAGCTTCAGGAAACTTTTGGAGTATGAAGTAGGTGGCCCTGTGTT 895
 QY 261 CysValGlnThrAlaTyrGlyValGluValGluAlaAsnProTyrAspProAsn 280
 DB 896 TGTGTCCAAACTGCTTATGTTGTGGAGTTCAGGAGGAGCAATCCATACGATCCCAAC 955
 QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300
 DB 956 TTAATGGATTTTATGAGCTACAGACTATATGCAACAGAAATTTACAGTGTGGAGAA 1015
 QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu 320
 DB 1016 GAATATCCAACATTTCTCTATGTCATGCCATGTGCGCAACAGACATTTTTTTTGAGAA 1075

QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer 340
 DB 1076 ACTGTTTGGCTCAAAAGATGCCATGCTTTCATCTACTGAAGAGAACTAATGTCA 1135
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTrpSerTyrIle 360
 DB 1136 CGATTGAAGACTCTGGTATCCAAAGTTACAAAATTTATGAAGAGAAATGGTCTTATTT 1195
 QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAla 380
 DB 1196 CTTGTTGGGGTCTTTACCAACACAGAGCAAAAGAACCTAGCATTTGGTGTGTCGACA 1255
 QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
 DB 1256 AGCATGTGATCCAGCAACAGCTTATCGTTGTAGCATCCTATCATCAGAGCTCCAAA 1315
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
 DB 1316 TATGCTTCTGTAATTCAAAAGATTTTGAAGCAAGATAACTCTGCATATCTGGTTCGGA 1375
 QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
 DB 1376 CAAAGCAGTGCAGTAAACATTTCAATGCAGCATGGAGCAGCTCTTGGCCAAAAGGCGA 1435
 QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
 DB 1436 AAACGTCAAGAGCATTTCTTCTTTCGGGTAGAGCTTATGTGCAGCTAGATATTGAA 1495
 QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480
 DB 1496 GCAACAGAACAGCTTCTTTAGAACCTTCTCGCTTGCACACTTGGATGTGGTGGGTTTC 1555
 QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 DB 1556 CTTGGGCTTCTCATCATCTTTCGATCTTTCGATCTTTCGATCTTTCGATCTTTCGATCT 1615
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 DB 1616 GCCCGCAACAGCATGAGGATGTCACITGTGAGACATTTGCTTTCAGATCTTCTGTCGA 1675
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 DB 1676 GTTATGTTAAAGCTTACCTCGAAGG 1702

RESULT 3
 US-09-323-998D-1
 ; Sequence 1, Application US/09323998D
 ; Patent No. US20020102631A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CUNNINGHAM JR., FRANCIS X.
 ; APPLICANT: SUN, ZAIREN
 ; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
 ; METHODS OF USE THEREOF
 ; FILE REFERENCE: 108172-09019
 ; CURRENT APPLICATION NUMBER: US/09/323,998D
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: 09/086,724
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 09/088,725
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 08/937,155
 ; PRIOR FILING DATE: 1997-09-25
 ; PRIOR APPLICATION NUMBER: 08/624,125
 ; PRIOR FILING DATE: 1996-03-29
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1860
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS

LOCATION: (109)..(1680)
US-09-323-998D-1

Alignment Scores:

Pred. No.: 2,19e-228 Length: 1860
Score: 1959.50 Matches: 376
Percent Similarity: 81.65% Conservative: 60
Best Local Similarity: 70.41% Mismatches: 81
Query Match: 71.44% Indels: 17
DB: 10 Gaps: 5

US-09-701-395A-23 (1-529) x US-09-323-998D-1 (1-1860)

```

QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTrp 16
Db 109 ATGGAGTGTGGTGGCGCTAGGAATTCACAGATTCGCGTATGCGTATG 168
QY 17 ThrPheGlyThrArgAsnLeuSerSerSerLysLeuAlaValTrpAsnLeuH 36
Db 169 AGTTGTGCGAAGAAATTCAGTGGTTAAGAGATACAGCTATAGGAATATTC 228
QY 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyCly-----SerGlySerArg 54
Db 229 ---TTGGT-----AGTGTACAGAGCTAGCGCGCGGAGAGTTCCGGTAGTGAG 273
QY 55 SerSerValAlaValLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
Db 274 AGTTGTGCGTGGTGGAGAGAGATTCGCTGACGAAGAGATTTGTGAAGCTGGTGGT 333
QY 75 SerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
Db 334 TCGAGATTCATTTGTTCAATGTCAGCAGACAAAGATGGATGAGTCAAGCTTAAGCTT 393
QY 95 AlaAspLysLeuProPheProPheGlyGluSerValMetAspLeuValValIleGly 114
Db 394 GTTGATAAGTTCCTTATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453
QY 115 CysGlyProAlaGlyLeuSerLeuAlaGluAlaAlaLysLeuGlyLeuLysValGly 134
Db 454 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
QY 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLys 154
Db 514 CTCATTGCTGCAGATCTCTCTTACTTAACTTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 573
QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaValSerValSerArgHisLeuLeuHis 194
Db 574 GATCTGGCTGCAAAATGTATTGAGCATGTTTGAGAGAGACTATTGTTATCTGAT 633
QY 175 AsnAspAlaProValLeuLeuGlyArgAlaValArgValSerArgHisLeuLeuHis 194
Db 634 GATGACAAAGCTATTACCTTGGCTGCTTATGGAAGAGTATGTCGAGCTTGGCTCCAT 693
QY 195 GluGluLeuLeuLysArgCysValGluSerGlyValSerValSerLysValGlu 214
Db 694 GAGAGCTTTTGGAGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCTGAGTGTCT 753
QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluPheIle 234
Db 754 AGCATAACAGAGCTCTGATGGCTTAGACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
QY 235 ProCysArgLeuAlaValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGlu 254
Db 814 CCCTCAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
QY 255 ValGlyGlyProArgValCysValGlnThrAlaValGlyValGluValGluAsn 274
Db 874 GTTGGTGGACCTAGAGTCTGTGCAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLys 294
Db 934 AGTCCATATGATCCAGATCAATGGTTTTCATGGATTACAGAGATTATACCAACAGAGAA 993

```

```

QY 295 LeuGlnCysSerGluGluGlyProThrPheLeuTyrValMetProMetSerProThr 314
Db 994 GTTCGGAGCTTGAAGCTGAGTATCCAGCTTCTGTGATCGCATGCTATGACAAAGTCA 1053
QY 315 ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
Db 1054 AGACTCTCTTCGAGAGACATGTTGGCTCAAAAGATGTCATGCCCTTGTGATTCCTA 1113
QY 335 LysArgLysLeuMetSerArgLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
Db 1114 AAACGAAGCTCATGTTAAGATTAGATACCTCGAATTCGATTTAAAGACTTACGAA 1173
QY 355 GluGluTrpSerTyrIleProValGlyCysLeuProAsnThrGluGlnLysAsnLeu 374
Db 1174 GAGGAGTGGTCTATATCCAGTGGTGGTTCCTTGCACAAACACGCAACAAAGATCTC 1233
QY 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394
Db 1234 CCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1293
QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer 414
Db 1294 TTGTCTCAAGCTCAAAATATGATCAGTCATCGCAGAGATCTTCTTCTTCTTCTTCTTCT 1353
QY 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetClnAlaTrpSerSer 434
Db 1354 AAACAGATCAACAGT-----AATATTCAGACAGCTTGGGATACT 1395
QY 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuLeuLeu 454
Db 1396 TTATGCGCCACCAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 455
QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPheArgLeuProThr 474
Db 1456 GTTCAATTCGATACCAAGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 514
Db 1516 TGGATGTGGCAGGTTTCTAGGATCAACATTAACATCAGGAGAGATCTGCTTCTTCTTCT 1575
QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
Db 1576 TTATCATGTTGGTCAATTCACCAACAAATTCAGAAAGAGTCTCATCAATCATCTCATC 1635
QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
Db 1636 TCTGATCCACCGGACCAACCATGATAAAACCTATCTCAAA 1677

```

RESULT 4

```

US-09-323-998D-24
; Sequence 24, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 24
; LENGTH: 1378
; TYPE: DNA

```

ORGANISM: Solanum tuberosum
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (7)..(11)
 OTHER INFORMATION: a, t, c, g, other or unknown
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (1180)..(1181)
 OTHER INFORMATION: a, t, c, g, other or unknown
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (1330)
 OTHER INFORMATION: a, t, c, g, other or unknown
 US-09-323-998D-24

Alignment Scores:
 Pred. No.: 5, 61e-170 Length: 1378
 Score: 1478.50 Matches: 281
 Percent Similarity: 86.77% Conservative: 47
 Best Local Similarity: 74.34% Mismatches: 48
 Query Match: 53.90% Indels: 3
 DB: 10 Gaps: 1

US-09-701-395a-23 (1-529) x US-09-323-998D-24 (1-1378)

QY 151 AspGluPheLysAspLeuGluArgCysIleGluHisAlaTrpLys-AspThrI1 170
 Db 14 GATGAGTCAAGATCTTGGTCTCAAGCCGTCATGAACATGTTGGCTGGGATACCAT 73
 QY 170 eValTyLeuAspAsnAlaProValLeuIleGlyArgAlaTyrgValSerAr 190
 Db 74 TGTATATCTTGATGATGATGATCTATCTTATTGGCGTCCATGGAAGATTAGTCG 133
 QY 190 qHisLeuHisGluGluLeuLysArgCysValGluSerGlyValSerTyLeuAs 210
 Db 134 CCATTACTGCAGAGATTACTCAAAAGGTGTGGAGCAGGTGTTTGTATCTAAA 193
 QY 210 pSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAs 230
 Db 194 CTCGAAAGTGATGAGTATTGTAGGCCACAAATGGCCACAGTCTGTAGAGTGGAGGG 253
 QY 230 nGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaSerGlyLysLe 250
 Db 254 TGATGTTGATTCCTTCGAGTTTGTGATGTTGTCATCGGGAGCAGCTCGGGGAAAT 313
 QY 250 uLeuGluTyrgLuuValGlyGlyProArgValCysValGlnThrAlaTyrgValGluVa 270
 Db 314 CTTCAGTATGAGTGGAGGTCTAGAGTTCTGTTCAACACGTTATGGAGTGGAGT 373
 QY 270 lGluValGluAsnAsnProTyrgAspProAsnLeuMetValPheMetAspTyrgAspTy 290
 Db 374 TGAGGTGATGAACAATCCATTGTACCGGAGCCTGATGTTTCATGGATTATAGACTA 433
 QY 290 rMetGlnGlnLysLeuGlnCysSerGluGluTyrgProThrPheLeuTyrgValMetPr 310
 Db 434 TGTGACAGACAGCTCAATCTTTAGAGGTAATAATCCCAATTTCTCTATGCCATGCC 493
 QY 310 oMetSerProThrArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetPr 330
 Db 494 CATGCTCTCAACACGAGTCTTTTCGAGGAACTTGTTCGCTTCAAAAGATGCAATGCC 553
 QY 330 oPheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValTh 350
 Db 554 ATTGCGATCTCTTAAGAAAAAATGATGTTACGATTGAACACCCCTCGGTGTAAGAAATAA 613
 QY 350 rLysValTyrgLuuGluGluTrpSerTyrgIleProValGlySerLeuProAsnThrGl 370
 Db 614 AGAANTTATGAGAGGAATGGTCTTATACATACCATGTTGGAGGATCTTTGGCCAAATACA 673
 QY 370 uGlnLysAsnLeuAlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrgSe 390
 Db 674 AAAAAAACACTTGATTTGCTGCTGCTAGCATGTTGATCATCCAGCCAGGTTTATTC 733

QY 390 rValValArgSerLeuSerGluAlaProLysTyrgAlaSerValIleAlaLysLeuLy 410
 Db 734 ACTCGTCAGATCACTGTCTGAAGCTCAAAATGCGCTTGGTCTTCAAAATATATACG 793
 QY 410 sGlnAspAsnSerAlaTyrgValValSerGlyGlnSerSerAlaValAlaAsnIleSerMetGl 430
 Db 794 ACATAATCATAGCAAGAAATATGTTACT---AGTTCAAGTACCCCGGATATTTCAACTCA 850
 QY 430 nAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGl 450
 Db 851 AGCTTGGAACTCTTTTGGCCACAGAACGAAACGAAAGATCGTTTTCCTTATTTGG 910
 QY 450 yLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePh 470
 Db 911 ACTGGCTCTGATATTGAGCTGGATATTGAGGGGATGAAGGTCAATTTTCGCGCGTTCCT 970
 QY 470 eArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLe 490
 Db 971 CCGTGTGCCAAAATGATGTGGCAGGAGTTCCTTGGTTCAGTCTTCTTA-CGAGACCT 1029
 QY 490 uValLeuPheSerMetTyrgMetPheValLeuAlaProAsnSerMetArgMetSerLeuVa 510
 Db 1030 CATGTTATTTGGCTTCTACATGTTTATTATGACCAAAATGACATGAGAGAGGCTTAAT 1089
 QY 510 lArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrgLeu 527
 Db 1090 CAGACATCTTTTATCTGATCTCTACTGTGCAACATGATGAAGACTTATCTT 1141

RESULT 5

US-10-041-472-3
 ; Sequence 3, Application US/10041472
 ; Patent No. US20020092039A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shemmaker, Christine
 ; TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALITY OILS
 ; TITLE OF INVENTION: SEEDS
 ; FILE REFERENCE: 16516.141
 ; CURRENT APPLICATION NUMBER: US/10/041,472
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 08/908758
 ; PRIOR FILING DATE: 1997-08-08
 ; PRIOR APPLICATION NUMBER: US 60/024145
 ; PRIOR FILING DATE: 1996-09-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1272
 ; TYPE: DNA
 ; ORGANISM: Brassica napus
 US-10-041-472-3

Alignment Scores:
 Pred. No.: 9, 8e-160 Length: 1272
 Score: 1394.00 Matches: 265
 Percent Similarity: 85.51% Conservative: 30
 Best Local Similarity: 76.81% Mismatches: 46
 Query Match: 50.82% Indels: 4
 DB: 12 Gaps: 2

US-09-701-395a-23 (1-529) x US-10-041-472-3 (1-1272)

QY 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
 Db 67 AGTGTGATGGATATCTGCAGAAATTCGGCTTGTGTTGGTCTCGTGGTGGCTGGCG 126
 QY 124 AlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr 143
 Db 127 GCTGAATCAGCTTAAGTTAGGACTTAAAGTTGGACTGATTGGTCTCGACCTTCTTCACT 186
 QY 144 AsnAsnTyrgLuuValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
 Db 187 AACAACTACGGTGTGGGAAAGATGAGTTCAACGATCTTGGCTTCAAAATATGATTGAG 246


```
Db 846 GGTTCCTGGCAACACGGAAGAAAGTCTCGCCTTTGGCGTGCAGCTAGCATGGTA 905
Qy 384 HisProAlaThr 387
Db 906 CATCCGCAACA 917

RESULT 7
US-10-041-472-4/c
; Sequence 4, Application US/10041472
; Patent No. US20020092039A1
; GENERAL INFORMATION:
; APPLICANT: Shemmaker, Christine
; TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALITY OILS IN SEEDS
; FILE OF INVENTION: SEEDS
; FILE REFERENCE: 16516.141
; CURRENT APPLICATION NUMBER: US/10/041.472
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 08/908758
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: US 60/024145
; PRIOR FILING DATE: 1996-09-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Brassica napus
US-10-041-472-4

Alignment Scores:
Pred. No.: 9, 41e-80 Length: 1590
Score: 739.00 Matches: 164
Percent Similarity: 56.70% Conservatives: 90
Best Local Similarity: 36.61% Mismatches: 144
Query Match: 26.94% Indels: 50
DB: 12 Gaps: 10

US-09-701-395a-23 (1-529) x US-10-041-472-4 (1-1590)
Qy 107 ValMetAspLeuValIleGlyCysGlyProAlaGlyLeuSerLeuAlaGluAla 126
Db 1304 GTTGTGATCTAGTATCTGTCGCGCGCGCGCTGGTGTAGCGGCTCAGCAAGTC 1245
Qy 127 AlaLysLeuGlyLysValGlyLeuIle-----GlyProAspLeuProPheThrAsn 144
Db 1244 TCCGAAGCTGACTCTCTGCTCTCTCCATCGATCCCTCCCAAACTCATTTGGCCTAAC 1185
Qy 145 AsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGluHis 164
Db 1184 AACTACGGAGTTTGGGTTGACGAGTTGCAAGCCATGACCTTGTAGACTGCCTCGACACC 1125
Qy 165 AlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArgAla 184
Db 1124 ACCTGTCCGCGCGCTGCTCTATCATCAGCAGCGCTCCGAAAAGGACCTGTCGCCGCCT 1065
Qy 185 TyrGlyArgValSerArgHisLeuHisGluLeuLeuLysArgCysValGluSer 204
Db 1064 TAGCGAGAGTGAACCTTAACAGCTCAATCCAGATGCTTCAGAACTGCATCACCAC 1005
Qy 205 GlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSer 224
Db 1004 GTGTGTAGATCCATCAGGCTAAGTCACTCAGCTGTTTCCAGGAG---GAGGTTAACTCC 948
Qy 225 LeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGly 244
Db 947 ACTGTTCTGTCAGTACGGGTGTGAAGATTCAGGCTCTCTGTTGTTCTCGACGCTACTGGT 888
Qy 245 AlaAlaSerGlyLysLeuLeuGluTyrGluValGly---GlyProArgValCysValGln 263
Db 887 ---TTTCAAGATCTTGGTTCAGTATGATAAGCCTTATTAACCTTGGG-----TATCAA 837
Qy 264 ThrAlaTyrGlyValGluValGluValGluValGluValGluValGluValGluVal 283
```

```
Db 836 GTAGCTTATGTTATCATCTGCTGAGTTCGATGTCACCGCTTGTGATGTGATAAGATGGTG 777
Qy 284 PheMetAspTyrArgAsp-----TyrMetGlnGlnLysLeuGlnCysSer 298
Db 776 TTCATGGAGTGGAGAGATAAGCATCTTGGCGGTACCCCTGAGCGTTAAAGCGGACACAGC 717
Qy 299 GluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePhe 318
Db 716 -----AAGATCCCTACGTTCTGTACGCGATGCGCTTTTCTTCCAAACAGGATCTTCTT 663
Qy 319 GluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLysArgLysLeu 338
Db 662 GAAGACAGCTCTCTGCTGCTGAGCGGTCTGAAGATGAAGATGATCAAGAGAGATG 603
Qy 339 MetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluIleTrpSer 358
Db 602 GTTGGCGAGCTGNAACATTTGGGGATCAACGTGAAGCGGATTCAGGAAGACGCGTGT 543
Qy 359 TyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAla 378
Db 542 GTGATCCCTATGGGAGGTCTCTTACCGGTCTTGCCTCAAGAGATGTTGTGGCATTTGCCG 483
Qy 379 AlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAla 398
Db 482 ACGCGCGGATGTTTCATCTCGACTGTTACATGTTCTAGGACTCTTGGACGCTCT 423
Qy 399 ProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValVal 418
Db 422 CCATAGTTGCAACGCTATAGTCCGG-----TACCTTGGT 387
Qy 419 SerGlyGlnSerSerAlaVal-----AsnIleSerMetGlnAlaTrpSerSerLeu 435
Db 386 GGTGGTAACAACAACGCGCTTGAGAGGAGATGAGCTCTCGCGTGGAGTGTGGAGAGACTTG 327
Qy 436 TrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuVal 455
Db 326 TGCCCTATTGAGAGCGGAGACAGAGGAGTCTCTCTGTTTGGGATGGATATTTGCTG 267
Qy 456 GlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrp 475
Db 266 AGCTTGTATTTGGATGCTACTAGGAGGTCTCTTGTATGCTTCTTGTACCTGGAACCGCGT 207
Qy 476 MetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuSerMet 495
Db 206 TACTGGCATGTTCTTCTGCTGCGAGGCTGTTCTCCCGGACCTGGTGTCTTCTCGGGCTG 147
Qy 496 TyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal-----510
Db 146 TCGCTCTCTCGCATGCTTCTATACCTCGAGGTGGAGATCATGACGAGGAAGAACTGTT 87
Qy 510 -----510
Db 86 CCTCTTGTAGATGATCAACAATTTGGTACAAGATAGAACCGAATTCACGACACTGG 27
Qy 511 ArgHisLeuSerAspProSer 518
Db 26 CGGCCGTTACTAGTGGTCCGAGC 3

RESULT 8
US-10-166-037-2
; Sequence 2, Application US/10166037
; Publication No. US20030087337A1
; GENERAL INFORMATION:
; APPLICANT: GIRAUD, Eric
; APPLICANT: HANNIBAL, Laure
; TITLE OF INVENTION: Isolated Carotenoid biosynthesis gene cluster involved
; in canthaxanthin production and application thereof
; FILE REFERENCE: 1721-55
; CURRENT APPLICATION NUMBER: US/10/166.037
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/297,272
; PRIOR FILING DATE: 2001-06-12
```


Thu May 22 09:39:17 2003

QY 109 AspleuValValIleGlyCysGlyProAlaGly-----LeuSerLeuAlaAlaGluAla 126
 Db 16 GATCTCATCTGCTGCGGCGGCTGCGCTATGCGCTTATGCGCTCGGCTTCAGCAA 75
 QY 127 AlaLeuLeuGlyValGlyValGlyLeuLeu-----GlyProAspLeuProPheThrAsn 144
 Db 76 CAGCATCCGGATATCGGATCTGCTTATGAGCGGCTCTGAG-----GCGGGA 126
 QY 145 AsnTyrGlyValTrpGluAspGluPheLeuAspLeuGluArg----- 160
 Db 127 GGAACCATACCTGCTTCCAGAGAGGATTTACGCTGAATCAGCATCCTCGGATA 186
 QY 161 -----CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaPro 178
 Db 187 GCGCGCTTGTGCTCCACTGCGCGACTACCGCTTCCGCTTCCGCTTCCGCTTCGC 246
 QY 179 ValLeuIleGlyArgAlaTyrGlyArgValSerArgHis-----LeuLeuHisGlu 195
 Db 247 CATGTGAACAGTGGCTACTACGCTGCGTCCGCTCCGCGGATTTCCGCGGATACCTCGGCAA 306
 QY 196 GluLeuLeuLysArgCys---ValGluSerGlyValSerTyrLeuAspSerLysValGlu 214
 Db 307 CAGTTGGACAAATTTATGCTGCATACCGCGTTTCAGCGCTTTCATGCTGAATCGGTC 366
 QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle 234
 Db 367 CAGTAGCGGATGCGCGGATTTATCATGCCAGTACAGTACGACGCGGCTTACACG 426
 QY 235 ProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuGluTyrGlu 254
 Db 427 COTGATTCGACCTACCGCTAGGATTCAGGCAATTTATCGGTACG-----GATGGCAA 480
 QY 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValAsn 274
 Db 481 CTGACGCGCGCG-----CATGTTTA-----TCG 504
 QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnLys 294
 Db 505 TCACCG-----ATTATCATGGAT----- 522
 QY 295 LeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThr 314
 Db 523 GCGACGCTGCATCAGCAAAATGCTACCGCTTTGTTTATACCTCGCGCTTCGCGCAAC 582
 QY 315 ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
 Db 583 GCACCTGCTGATGAGACACACACTACATTTGACAGGCTTAATCTTCAGCGCGGCGG 642
 QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
 Db 643 CGTCAGAACATTCGCGGATGCTGCGGACAGGGTGGCGGTACAGAGCTTGTGCGG 702
 QY 355 GluGluTrpSerTyrIleProValGlyGlySerLeuProAsn-----ThrGlu 370
 Db 703 GAGACAGCGGTGCTATGCTGCTTACGTTAACGCGGATTAATCGTCAGTTTGGCAACAG 762
 QY 371 GlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSer 390
 Db 763 CAACCGCAACGCTGTAGCGGATTCGCGCGGCTGTTTATCCGACACCGGCTACTCC 822
 QY 391 ValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLys 410
 Db 823 CTACCGCTGCGGTGGCGCTGGCGATCGTCTC----- 855
 QY 411 GlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAla-----ValAsn 426
 Db 856 -----AGCGGCTGGATGCTTTTACCTCTTCTCTCTTCCAGCAGCATGCTGCTC 906
 QY 427 IleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhe 446
 Db 907 TTTGCCAGCAACGCTGG-----CAGCAACAGGGGTTT 939

QY 447 Phe---LeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhe 465
 Db 940 TTCCGCATGCTGAATCGCATGTTGTTTACCGCGACCGCGGCTGCGCTG 999
 QY 466 PheArgTrpPhePheArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSerLeu 485
 Db 1000 ATGCAGCGTTCCTATGCTTACCGGAGGATTTGATTCGCCGCTTTTATGCGGGAAACTC 1059
 QY 486 SerSerPheAsp 489
 Db 1060 ACCGTGACCGAT 1071
 RESULT 10
 US-09-938-842A-1578
 ; Sequence 1578, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1578
 ; LENGTH: 1593
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1578
 Alignment Scores:
 Pred No.: 0.000112 Length: 1593
 Score: 123.00 Matches: 119
 Percent Similarity: 36.30% Conservative: 77
 Best Local Similarity: 22.04% Mismatches: 182
 Query Match: 4.48% Indels: 163
 DB: 9 Gaps: 27
 US-09-701-395A-23 (1-529) x US-09-938-842A-1578 (1-1593)
 QY 77 LeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAsp 96
 Db 73 CTCTTATACGTCCTCGTCGCGGAGCAGACGATTCATGCTGTCATGTCGTAAC 132
 QY 97 LysLeuProIleProPheGlyGluSerValMetAspLeuValIleGlyCysGly 116
 Db 133 GGAACCCCTAACGCTGAATCTGGAACACGCTT---GATATTATCATGTCGCTGCT 189
 QY 117 ProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGlyLeuIle 136
 Db 190 GTCGCGGCGCTGCGCTTCTCATACCTCGGCAAGGAGGAGAGAGTTCACGTTATA 249
 QY 137 GlyProAspLeuProPheThrAsn----- 144
 Db 250 GAAAGAGACTTACGGAGCCTGATCGAATTCGCTGGAATTAATTCAGCTGCTGTTAC 309
 QY 145 -----AsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuArg 160
 Db 310 TTCAAGTTAATCAACTCGGCTT---GAAGATTGTTGTGAAGGATATAGATCGCGAGACA 366
 QY 161 CysIleGluHisAla---TrpLysAsp-----ThrIleValTyr---LeuAsp 174
 Db 367 GTTCTTGGTTATGCTCTCTTTAAAGATGGGAACACACTACTCTTACCGTTGGAT 426


```

QY 175 AsnAspAlaProValLeuIleGlyArgAlaTyr-----GlyArgValSerArgHisLeu 192
Db 427 CAGTTTGATTCGAGTTCGGGGTCTAGCTTTCAATAGGAGATTTGTGAGAGGATG 486
QY 193 LeuHisGluGluLeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLys 212
Db 487 CGAGAGAAAGCTCTTTACTTCCCAATGTTCAATGGAGCAAGCAAGTGCATCGTGTG 546
QY 213 ValGluArgIleThrGluAlaGly-----AspGly 222
Db 547 GTGGAA-----GAAACGGAATATCAAAAGGTGTTCAATACAAACCAAGATGCG 597
QY 223 HisSerLeuValValCysGluAsnGluIlePheProCysArgLeuAlaThrValAla 242
Db 598 CAAGAGCTT-----AAGTCATTTGCTCT-----CTCAGTATTGTGTGT 636
QY 243 SerGlyAlaAlaSerGlyLysLeuLeuLysGluValGlyGlyProArgValCysVal 262
Db 637 GATGTTGTTTCTCAAC-----TTGCGTCGATCTCTCTCAAAACCTAAGGTGTGT--- 687
QY 263 GlnThrAlaTyrGlyValGluValGluValGluLysLeuGlnCysSerGluGluLutyr 302
Db 688 -----AACATGGTTGTGGAGTGCATCCAT----- 714
QY 283 ValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluLutyr 302
Db 715 -----TTGTTGGTCTGTTATGAGAAATTCGCAACTCCCGTTT 753
QY 303 ProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePhe----- 318
Db 754 CCGAACACGCGGACGTTGTTGCGTCGATCTCTGCAATTTATTCATCCTATCAGC 813
QY 319 -----GluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArg 336
Db 814 AGCTCGAAGTCCGTTGCTTGGTA-----GATGTACCAGTTCA 852
QY 337 LysLeuMetSer-----ArgLeuLysThrLeu-----GlyIle 347
Db 853 AAATCTCTTCAGTTCGAGTGGTGAATGGCTCACCATCTCAAAACAATGTTGACCG 912
QY 348 GlnValThrLysValTyrGluGluLutyr-----SerTyrIleProValGlyGly----- 364
Db 913 CAGGTACCACTCAGATCCGTCGTGATGCTTCTTCGTCAGTGCAGAAAGGTAACATAGA 972
QY 365 SerLeuProAsnThrLutyr-----AsnLeuAlaPhe 376
Db 973 ACATGCCGCAAGAGCAAGCAAGTGCATGATCAATTCATACACCTGGAGCTTTGCTTTA 1032
QY 377 GlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSer 396
Db 1033 GGTGATGCTTCAACATCGCCATCCTCTTACTGGAGTGGTATGACCGTTGCATGTCT 1092
QY 397 GluAlaProLysTyrAlaSerValIleAlaLysIleLeu-----LysGlnAsp 412
Db 1093 GATATAGTATCCTCCGTGATCTATTGAACCGCGTCGTCGATTAACCAACAAGAGTCC 1152
QY 413 AsnSerAlaTyrValValSerGlyClnSerSerAlaValAsnIleSerMetGlnAlaTyr 432
Db 1153 TTATCCAAATACATAGATCA----- 1173
QY 433 SerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPhe-GlyLeuG1 452
Db 1174 -----TTCTACACATTCGGGAAACCG 1194
QY 452 uLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePheArgLe 472
Db 1195 GTTGCTT---CGATATCAATACATACGCGCGCTCTTTACAAAGTCCTTTTAGCATCT 1251
QY 472 uProThrTrp----- 475
Db 1252 CCAGACATGCAAGACGAATGCGTCGAGCTTGCTTGCATTTATCTTACCTCGGAGG 1311
QY 476 -----MetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLe 490

```

```

Db 1312 GTTTCCTCATCAGACCTGTGCTCTTGTCTATCT-GGTTGAACCCAGACCTATGAGCCT 1370
QY 490 uValLeu-----PheSerMetTyrMetPhe-----ValLeuAlaPr 502
Db 1371 TGTTCCTCATCTTCGAGTTGCGATTTCGGGGTGGTTCGTTGCTTGTACCTCTCCC 1430
QY 502 oAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAlaVal 521
Db 1431 GTCCGTTAAACGGTTATGGCTTGAGCTAGACTAATCTCGATGCTTCAGGGATCATA 1488

```

RESULT 11

```

US-09-738-626-529
; Sequence 529, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 529
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-529

```

```

Alignment Scores:
Pred. No.: 0.000636 Length: 1272
Score: 115.50 Matches: 76
Percent Similarity: 35.69% Conservative: 55
Best Local Similarity: 20.71% Mismatches: 137
Query Match: 4.21% Indels: 99
DB: 9 Gaps: 16

```

US-09-701-395A-23 (1-529) x US-09-738-626-529 (1-1272)

```

QY 103 PheGlyGluSerValMetAspLeuValValIleCysGlyProAlaGlyLeuSerLeu 122
Db 7 TTTGTGTCACAACTTTTGTGATGTTGATCATCGCGCGGCCCTCAGTGCCAGCGCC 66
QY 123 AlaAlaGluAlaLysLeuGlyLeu----- 131
Db 67 GCGGTCCATCGCGCCAGGACTGGGCTTCAACACATTGCTTATCGACCGCTCTCTCCCG 126
QY 132 -----LysValGlyLeuIle 136
Db 127 CGGGATAAACGTCGTGGCGATGGCTTACTCCCGTCGCGATTCACCGAGCTAGACTTCTA 186
QY 137 GlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspClnPheLysAspLeu 156
Db 187 GGTGTGTCGATCAGGTATACCGGGGATTTAT-----TTCAACAAGGGCTTGAACACTGCAT 240
QY 157 GlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTrpLeuAspAsnAsp 176
Db 241 GGTGTTGGTGTCTGTTTCAGCGCGCGGAGACA-----TATTTCAGGAATAG 294

```


us-09-701-395a-23.rnpb

Thu May 22 09:39:17 2003

COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patencin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/547,267
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: RAN 6002/170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-547-267-8

177 AlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGlu 196
295 -----GGTCCCGATGTCGCGATGGAG-----TTCCGATGAT 327
197 LeuLeuLysArgCysValGluSerGlyVal----- 206
328 TTGTTATTCGCGTGGCAAAATCTCATGAGGAAGTAACACCGTGGAGAACCGGAGCC 387
207 -----SerTyrLeuAspSerLysValGluArgIleThrGluAla 219
388 CAAGACCCCATTTTGGAGGGGAATTTCTTGGAGGCGTTGT-----ATTAATCAAGCA 441
220 GlyAspGlyHisSerLeuValValCysGluAsnGluIlePheLeuProCysArgLeuAla 239
442 GGG-----CAAGAGAAACCCCTCAAGCGGAGCAGATG 474
240 ThrValAlaSerGlyAlaAlaSer-----GlyLysLeuLeuGluTyrGluValGlyGly 257
475 ATTATTCGCGATGTCGCGTCCCTTCCCTTCCGTAAGAACTG-----GGT 519
258 ProArgValCysValGlnThrAlaTyrGlyVal-----GluValGlu 271
520 AGCAGTGGCAACCGATGAGGTATGATGCGCTGCTGCTTATTGTGAACCTCG 579
272 ValGluAsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMet 291
580 CTGCTGATGAACCGTGGATTCACCTCCCATGTGGAACTGCCGATGAAGATGGTGTGTG 639
292 GlnGlnLysLeuGlnCysSerGluGluTyrProThrPheLeuTyrValMetProMet 311
640 CAG-----CCAGGATATGGTGGATTTTCCCGCTG 669
312 SerProThrArgLeuPhePheGluGluThrCysLeuAlaSerLysAlaMetProPhe 331
670 GGCACGGCAGCGGTAATTTGGTGTGGCGGCTC-----TCACGGATACGAGACCG 726
332 AspLeuLeuLysArgLysLeuMetSer-----ArgLeuLysThrLeuGly 346
727 AGATCAATACGAGAAATTTGTAGCTTCTATGCGGGTCAGCGTCGTAAGCATGG--- 783
347 IleGlnValThrLysValTyrGluGluGluTyrPheTyrIleProValGlySerLeu 366
784 ---CAACTCGGCGCCGAGCAGCAGCTGCGCTTCCCTGCTGCTTATGGCGCGGGTG 840
367 ProAsnThrGluGlnLysAsn---LeuAlaPheGlyAlaAlaAlaSerMetValHisPro 385
841 TCGAATGTCGCGCGGCAACTGGATGCTGATCGCGGATTCGCGCGGTGTGTGAACCG 900
386 AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIle 405
901 CTGAACGGCAAGGCATCGACTATGCGCTGGAAACCGCGCGGATGCGCGTTCGACAGCTT 960
406 AlaLysIleLeuLysGlnAsp 412
961 GTGAAACCCCAAGCGCAT 981

RESULT 12
US-09-547-267-8
Sequence 8, Application US/09547267
Patent No. US20020147371A1

GENERAL INFORMATION:
APPLICANT: Hohmann, Hans-Peter
APPLICANT: Pasamontes, Luis
APPLICANT: Tessier, Michel
APPLICANT: van Loon, Adolphus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ

Alignment Scores: 1149
Prod. No.: 83
Score: 113.00
Percent Similarity: 33.82%
Best Local Similarity: 20.34%
Query Match: 118
DB: 18

US-09-701-395a-23 (1-529) x US-09-547-267-8 (1-1149)

QY 109 AspLeuValValIleGlyCysGlyProAlaGly-----LeuSerLeuAlaAlaGluAla 126
Db 10 GATCTGTGATCGCGGCGGCGGCTGCGGTGCGGTGATCGCGTTCGCGTTCGCGGAC 69
QY 127 AlaLysLeuGlyLeuLysValGlyLeuIle-----GlyPro 138
Db 70 GCGACCGGATGCGCGCATGCTGATGCGCGCGGCTCGCGGCTCGGACGAC 129
QY 139 -----AspLeuProPheThrAsnAsnTyr----- 146
Db 130 ACCTGGTCTCGCACGACGATCTTTCGCCGATGCTGCGCGCTGTCGCCCAT 189
QY 147 -----GlyValTrpGluAspGlu-----PheLysAspLeuGlyLeuGluArgCys 161
Db 190 CGTCGCGGGAATGACGAGATCGAGGATCGCGTTCGCCGAC----- 231
QY 162 IleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIle 181
Db 232 -----CATTCGCGCGCTCGACGAGGATGCTGATGCGCGGCTCGCGGCTCGCGT 285
QY 182 GlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLysArgCys 201
Db 286 GGGCTCTCGAGGTGTCGATCTCGGTGGATACGAT----- 324
QY 202 ValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAsp 221
Db 325 -----GTCGCGAGCTGGACGATACCGCGCGGCGGCGGCTCGCGG 372
QY 222 GlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrVal 241
Db 373 ATCGAGGCTCGCGTGTGATCGAC----- 396
QY 242 AlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCys 261

D	b		397	GCCGCTGGTGC----- ---GTGAGACCGGCCACCTGAACC	429
O	y		262	Val-----GlnThrAlaTygIyValGIuValGIuValAsnAsnProTyArgSPpro	279
D	b		430	GTGGGTTTCAGAAATTCGTGGCGCTCGAGATCGAGACCGAC--GCCCCCCATGGCGTC	486
O	y		280	AsnLeuMetValPheMetasp::: ----TyArgAspTyMet	291
D	b		487	GAGCGCCGATGATCATGACGCGACCGCTCCGCAGATGACGGGTACCGC-----	537
O	y		292	GlnGlnLysLeuGlnCysSerGIuGluGluTyRProThrPheLeuTyValMetProMet	311
D	b		538	-----TTCATCTACTGCTGCCCTTC	558
O	y		312	SerProThrArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPhe	331
D	b		559	AGTCCCACCGCATCTGATCGAGATACGGCTACGACGACGGCGCGGATCTGCACGAT	618
O	y		332	AspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLys	351
D	b		619	GGCGCGCTGGCGAGCGGTCTGGACTATCCGCCAGCGGGGTGGACGGCGGAGG	678
O	y		352	ValTyRGlulGluGluTrpSerTyRileproValgly-----Glyser	365
D	b		679	ATCGCGCGCAAGG---GGCATCTCCCATCGCGCTGGCCCATGACGCCATAGGCTC	735
O	y		366	LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro	385
D	b		736	TGGCGGCGCACCGCGCACGGGGCGGTGCCGGTGTTGGCTGGGGCGAGGGCTTCCACCCC	795
O	y		386	AlaThrGlyTyRSerValValArgSerLeuSerGluAlaProLysTyRAlaSerValIle	405
D	b		796	GTACCGGATATTCGTG-----CCCTATGCGCGCGAGGTCCG	834
O	y		406	AlaLysIleLeuLysGlnAspAsnSerAlaTyRValValSerGlyGlnSerSerAlaVal	425
D	b		835	GATGCCATCTGGCGCGCGCACCTGACG-----ACCGCGTCCGCG	873
O	y		426	AsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAla	445
D	b		874	CGTGCGGGTTCGCGGTGGGCCATC-----GATCGCGGGATCGGACCGCTC	924
O	y		446	PhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhe	465
D	b		925	CTGGCGGTCTGAACCGGATGCTGTTCGCGGGTCCCGCCGACCGCTGATGCGCTG	984
O	y		466	PheArgThrPhePheArgLeuPro	473
D	b		985	CTGACGGGTTCTACCGCTGCGG	1008

RESULT 13

```

US-09-920-923-1/c
; Sequence 1, Application US/09920923
; Publication No. US2003002273A1
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/09/920,923
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8625

```

TYPE: DNA
ORGANISM: Flavobacterium sp. R1534
FEATURE:
NAME/KEY: unsure
LOCATION: (8348)..(8349)

[illegible]

QY 139 -----AspLeuProPheThrAsnAsnTyr----- 146
Db 7495 ACCTGTCCTGCACACAGCATTTTCGCCCGAATGCTGGCGCCTGTCCGCCATT 7436
QY 147 -----GlyValTrpGluAspGlu-----PheLysAspLeuGlyLeuGluArgCys 161
Db 7435 GTCGCGCGGAATGGAGCATGAGGAGTGCGGTTCGCGAC----- 7394
QY 162 ileGluHisAlatrplysaspThrIleValTyrLeuAspAsnAspAlaProValLeuile 181
Db 7393 -----CATTCGGCGCCCTGCAGCAGCAGCTATGCTGCATGAGGCGGCGCGCTGATC 7340
QY 182 GlyArgAlatryrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArgCys 201
Db 7339 GGGTGTGTCAGGAGTGTGCATCTCGGTGGAATACGCA----- 7301
QY 202 ValGluSerGlyValSerTyrLeuAspSerylsValGluArgIleThrGluAlaGlyAsp 221
Db 7300 -----GTCGCGACCTGGACGATACCGCGCGAGCTGACGAGCGCTCGCG 7253
QY 222 GlyHisSerLeuValCysGluAsnGluIlePheIleProCysArgLeuAlaThrVal 241
Db 7252 ATCGAGGTCCTCGCTGATCGAC----- 7229
QY 242 AlaSerGlyAlaAlaSerGlyLysLeuLeuGluTrpGluValGlyGlyProArgValCys 261
Db 7228 GCCGCTGGTCCC-----GTCGAGACCCCGCACCTGACC 7196
QY 262 Val-----GlnThrAlatryrGlyValGluValGluValGluValGluValMet 291
Db 7195 GTGGGTTTCAGAAATTCGTGGCGCTCGAGATCGAGACCGAC-----GCCCCCATCGGCTC 7139
QY 280 AsnLeuMetValPheMetAsp-----TyrArgAspTyrMet 291
Db 7138 GAGCGCCCGATGATCATGAGCGACCGCTCCGACATGAGCGGTACCGC----- 7088
QY 292 GlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProMet 311
Db 7087 -----TTCATCTATCTGTGCGCTTC 7067
QY 312 SerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetProPhe 331
Db 7066 AGTCCACCCCGCATCTCTGATCAGATACCGCTACGCGACGCGCGCATGTGGACGAT 7007
QY 332 AspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlylleGlnValThrLys 351
Db 7006 GCGCGCTGCGCAGCGCTCGCTGGACTATCCCGCAGCGGGGCTGACCGGCGAGG 6947
QY 352 ValTyrGluGluGluTrpSerTyrIleProValGly-----GlySer 365
Db 6946 ATCGCGCGCGAAAGG---GGCATCTCGTCCCATCGCTGCGCCCATGACGCGCATAGGCTTC 6890
QY 366 LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro 385
Db 6889 TGGCGCGACACCGCGCGGGGGCGGTGGCTGGCTGGCGGAGGCTGTTCACACCC 6830
QY 386 AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValile 405
Db 6829 GTCACCGGATATTCGCTG-----CCCTATCGCGCGCGAGTGC 6791
QY 406 AlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaVal 425
Db 6790 GATGCCATFCGCGCGCGCGGACCTGACG-----ACCGCGTCCGCC 6752
QY 426 AsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAla 445
Db 6751 CGTCGCGGCTGCGCGCTGGCGCATC-----GATCGCGGATCGGACCGCTTC 6701
QY 446 PhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhe 465
Db 6700 CTGCGCGCTGACCGCATGCTGTTCGCGGCTGCCCGCGCGGCTGCTATCCCTG 6641

RESULT 14
US-09-920-923/c
Sequence 27, Application US/09920923
Publication No. US2003002273A1
GENERAL INFORMATION:
APPLICANT: Pasamontes, Luis
TITLE OF INVENTION: Improved Fermentative Carotenoid Production
FILE REFERENCE: Improved Fermentative Carotenoid
CURRENT APPLICATION NUMBER: US/09/920,923
PRIOR FILING DATE: 2001-08-02
PRIOR FILING DATE: 1997-12-01
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 27
LENGTH: 11233
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Plasmid pze4
US-09-920-923-27

Alignment Scores:
Pred. No.: Length: 11233
Score: 0.0435
Matches: 83
Conservative: 55
Percent Similarity: 33.82%
Best Local Similarity: 20.34%
Query Match: 4.12%
DB: Indels: 118
Gaps: 18

US-09-701-395A-23 (1-529) x US-09-920-923-27 (1-11233)
QY 109 AspLeuValValIleGlyCysGlyProAlaGly-----LeuSerLeuAlaGluAla 126
Db 7615 GATCTGCTGATCGCGGCGCGGGCTGTCGGTGGCTGATCGCGCTGCGGCTGCGCGAC 7556
-----Glypro----- 138

RESULT 14

US-09-920-923-27/c
US-09-920-923-27/c

; sequence 27, Application US/0322032A1
; sequence 28, Application NO US20030022273A1

PUBLICATION NO: 00000000
GENERAL INFORMATION:

APPLICANT: Pasamontes, Luis

APPLICANT: Tsygankov, Yuri

TITLE OF INVENTION: Improved Fermentive

FILE REFERENCE: IMPROVED FEMINIST;
PUBLICATION NUMBER: US/09/9

; CURRENT APPLICATION NUMBER: 2001-08-02
 ; CURRENT FILING DATE: 2001-08-02

;; CURRENT FILING DATE: 08/19/2010
;; PRIOR APPLICATION NUMBER: 08/980,833

PRIOR FILING DATE: 1997-12-01

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 27
; SEQ ID NO 28

```

;
; LENGTH: 11233
; TYPE: DNA

```

ORGANISM: Unknown

FEATURE:

[illegible]

US-09-920-923-27

Alignment Scores:

Alignment Scores:	Pred. No.:
0.0435	

score: 113.00

Percent Similarity: 33.82%

Best Local Similarity: 20.34%
4.72%

Query Match:

DB:

US-09-701-395A-23 (1-529) x US-09-920

1000

QY
109 AspLeuValValIleGlyCysGlyPhe

[illegible]

Dd 7615 GATCTGCTGATCGCCGCCGCCGCCTT

127 AlaValLeuGlyLeuLysValGlyLeu

Qy	14/11/2006	11
		11

Db 7555 CGCAGACCGGATGCGCGCATCGTGAT

1

1

QY 466 PheArgThrPhePheArgLeuPro 473
 Db 6640 CTGACGCGGTCTACCGCTGCCG 6617
 RESULT 15
 US-09-939-964-1/c
 ; Sequence 1, Application US/09939964
 ; Publication No. US20030054522A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosenthal, Andre
 ; APPLICANT: Freiberg, Christoph
 ; APPLICANT: Perret, Xavier Philippe
 ; APPLICANT: Broughton, William John
 ; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
 ; FILE REFERENCE: CARP0068
 ; CURRENT APPLICATION NUMBER: US/09/939,964
 ; PRIOR FILING DATE: 2001-08-27
 ; PRIOR FILING DATE: 09/214,808
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 536165
 ; TYPE: DNA
 ; ORGANISM: Rhizobium
 US-09-939-964-1
 Alignment Scores:
 Pred. No.: 34.5 Length: 536165
 Score: 111.50
 Percent Similarity: 41.28%
 Best Local Similarity: 22.55%
 Query Match: 4.06%
 DB: 9
 Gaps: 9
 US-09-701-395a-23 (1-529) x US-09-939-964-1 (1-536165)
 QY 54 ArgSerSerValAlaTyrLysGluGly-PheValAspGluGluAspHe-----IleY 71
 Db 295477 CGACGATCTATCGATTACCGCGGTGTTTTCAGGAGGACTTCTACGGCCCTGATAA 295418
 QY 71 sAlaGlySerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysG 91
 Db 295417 CAAGGCAACGGTGTCTTCTCTGTCAGTGGCGGTGCTGCGCGCGCGCGGACA 295358
 QY 91 nAlaLysLeu-----AlaAspLysLeuProIleProPheG 104
 Db 295357 GGTTTCCTTGGTCGCGGATACCGTCTGCGAGCGATATGCTCTCGGACTGCCTGCGC 295298
 QY 104 yGluSerValMetAsp-----LeuValIleGlyCysGlyProAlaGlyLeuSerLe 122
 Db 295297 GAGTGTCTTGAAGCAAGAAAGGTTGTTCTGAGGTGCGGCGCATCGATTCGATCTCGT 295238
 QY 122 uAlaAlaGluAlaAlaLysLeuGlyLeu---LysValGlyLeuIleGlyProAspLeuPr 141
 Db 295237 TGCAGTCGAGCTGTCGCGCTCAGCGCTGCGTACGTCAGTCATCATTCGATCTCGT 295178
 QY 141 oPheThrAsnAsnTyrGlyValTyrGluAspGluPheLysAspLeuGlyLeuGluArgCy 161
 Db 295177 CGAGCCTGGCAACACGGTCACATGG-----CGCGTGGTCTGTC 295139
 QY 161 stleGluHisAlaTrpLysAspThrIleVal-----TyrLeuAspAsnAlaPro-- 178
 Db 295138 GGTGTGGGGCGCTGCAAGACACCGCGGTCGACGATTTCCTATCACTACCTCTG 295079
 QY 179 -ValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLe 198
 Db 295078 GACAAATGTGCGGCGTGGACACCGCTAAAGGGCTCAGCAATCTCGAATGTGGACGAGT 295019
 QY 198 uLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGl 218
 Db 295018 CCGCAACTCGAGGGAATCCCATGGTTGCTTCGACATTCATCGAGGACCGCGACAT 294959

QY 218 uAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArg-- 237
 Db 294958 TGTGGTAGACACTTCGGCTCGACCGAGTGTACAGGGCGCTTGGCCTACATGTCCCGTAG 294899
 QY 238 -----LeuAlaThrValAlaSerGlyAlaAlaSerGlyAlaAlaSerGlyLysLeuLe 251
 Db 294898 CATAGGAAAGCGGTACTCTCGGCCACGCAACGAGCGCGCGGCGCGGCGCTTGGC 294839
 QY 251 uGluTyrGluValGlyGlyPro-----ArgValCysValGln 263
 Db 294838 GAGATTCAAACCGGAGCGCGGTGTTATGTTCTGCTGCTGAG 294796

Search completed: May 22, 2003, 01:22:38
 Job time : 573 secs

us-09-701-395a-23.rni

Thu May 22 09:39:17 2003

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:18:48 ; Search time 68 Seconds
(without alignments)
2385.765 Million cell updates/sec

Title: us-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRNLSSCPVWTFCT.....VRHLLSDPSGVWVAYLER 529

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 88274

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framet_p2n.model
-Q=/cgn2_1/USPTO_spool/US09701395/runat_15052003.165752.3396/app_query.fasta_1.711
-DB=Issued Patents NA -OFT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09701395 -CGN1_1_40 -runat_15052003.165752.3396 -NCPUS=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1959.5	71.4	1860	1	US-08-624-125-1 Sequence 1, Appl
2	1934	70.5	1887	4	US-09-201-641-5 Sequence 5, Appl
3	1394	50.8	1272	4	US-08-908-758-3 Sequence 3, Appl
4	1202	43.8	1960	4	US-08-908-758-2 Sequence 2, Appl
5	750.5	27.4	1650	1	US-08-399-561-5 Sequence 5, Appl
6	750	27.3	1959	4	US-09-201-641-1 Sequence 4, Appl
7	739	26.9	1590	4	US-08-908-758-4 Sequence 4, Appl
8	737	26.9	1614	1	US-08-399-561-4 Sequence 11, Appl
9	722.5	26.3	2897	4	US-09-134-607A-11 Sequence 15, Appl
10	722.5	26.3	3265	4	US-09-134-607A-15 Sequence 8, Appl
11	719.5	26.2	1666	4	US-09-134-607A-8 Sequence 13, Appl
12	719.5	26.2	1666	4	US-09-134-607A-13 Sequence 1, Appl

13	719.5	26.2	2876	4	US-09-134-607A-9 Sequence 9, Appl
14	719.5	26.2	2876	4	US-09-134-607A-14 Sequence 14, Appl
15	705.5	25.7	1756	2	US-08-702-588-1 Sequence 1, Appl
16	703.5	25.6	1740	4	US-09-134-607A-10 Sequence 10, Appl
17	703.5	25.6	1740	4	US-09-134-607A-12 Sequence 12, Appl
18	703.5	25.6	1740	4	US-09-134-607A-16 Sequence 16, Appl
19	606	22.1	4928	1	US-08-399-561-1 Sequence 1, Appl
20	137.5	5.0	1235	1	US-08-095-726-13 Sequence 13, Appl
21	137.5	5.0	1235	1	US-08-096-623A-13 Sequence 13, Appl
22	137.5	5.0	1893	3	US-09-147-009-5 Sequence 5, Appl
23	134.5	4.9	1235	1	US-08-095-726-15 Sequence 15, Appl
24	134.5	4.9	1235	1	US-08-096-623A-15 Sequence 5, Appl
25	129	4.7	1161	1	US-08-663-310-5 Sequence 5, Appl
26	129	4.7	1161	2	US-09-006-491-5 Sequence 5, Appl
27	129	4.7	1161	3	US-09-335-919-5 Sequence 7, Appl
28	129	4.7	2886	1	US-08-663-310-7 Sequence 7, Appl
29	129	4.7	2886	2	US-09-006-491-7 Sequence 7, Appl
30	129	4.7	2886	3	US-09-335-919-7 Sequence 9, Appl
31	125.5	4.6	1149	1	US-07-783-705A-9 Sequence 13, Appl
32	125.5	4.6	6918	1	US-07-783-705A-13 Sequence 8, Appl
33	113	4.1	1149	3	US-08-660-645A-8 Sequence 8, Appl
34	113	4.1	1149	3	US-09-298-718-8 Sequence 8, Appl
35	113	4.1	1149	4	US-09-546-969-8 Sequence 1, Appl
36	113	4.1	8625	4	US-08-980-832-1 Sequence 27, Appl
37	113	4.1	12333	4	US-08-980-832-27 Sequence 3, Appl
38	112	4.1	1748	3	US-09-147-009-3 Sequence 11, Appl
39	109.5	4.0	1650	2	US-08-500-635A-11 Sequence 1, Appl
40	109.5	4.0	1650	4	US-09-167-151-11 Sequence 1, Appl
41	109.5	4.0	4656	4	US-09-425-665-1 Sequence 1, Appl
42	109.5	4.0	4656	4	US-09-685-668-1 Sequence 1, Appl
43	104.5	3.8	1756	3	US-09-147-009-1 Sequence 1, Appl
44	99.5	3.6	2443	2	US-08-745-934-2 Sequence 2, Appl
45	99.5	3.6	11340	4	US-08-961-527-147 Sequence 147, App

ALIGNMENTS

RESULT 1
US-08-624-125-1
; Sequence 1, Application US/08624125

; Patent No. 574341

; GENERAL INFORMATION:

; APPLICANT: CUNNINGHAM JR., FRANCIS X.

; APPLICANT: SUN, ZAIREN

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND

; TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

; ADDRESSEE: P.C.

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/624,125

; FILING DATE: 29-MAR-1996

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: KELBER, STEVEN B.

; REGISTRATION NUMBER: 30,073

; REFERENCE/DOCKET NUMBER: 2747-063-27

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1860 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 109..1680
 ; OTHER INFORMATION: /product= "E-CYCLASE FROM A.
 ; OTHER INFORMATION: THALIANA"
 ; US-08-624-125-1

Alignment Scores:

Pred. No.:	3,77e-220	Length:	1860
Score:	1959.50	Matches:	376
Percent Similarity:	81.65%	Conservative:	60
Best Local Similarity:	70.41%	Mismatches:	81
Query Match:	71.44%	Indels:	17
DB:	1	Gaps:	5

US-09-701-395a-23 (1-529) x US-08-624-125-1 (1-1860)

```

QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTrp 16
DB 109 ATGGAGTGTGGGGCTAGGAATTCGACGAATGGCGGTTTCAACATTTCCGTCATGG 168

QY 17 ThrPheGlyThrArgAsnLeuSerSerSerLeuAlaTyrAsnIleHisArgTyrGly 36
DB 169 AGTTGCGAAGGAATTTCCACGTGTTAAGAGATACAGCTATAGGAATATTCGTTTCGGT 228

QY 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerArg 54
DB 229 ---TTGTGT-----AGTCTCAGCTAGCGCGCGGGAAGTTCCGGTAGTGAG 273

QY 55 SerSerValAlaTyrLysGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
DB 274 AGTTGTAGCGGTGAGAGAAGATTTCGCTGACGAGAGAGATTTCGTAAGCTGGTGT 333

QY 75 SerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
DB 334 TCTGAGATTCTATTGTTCAAAATGCAGCAGCAACAAAGATATGGATGAACAGTCAAGCTT 393

QY 95 AlaAspLysLeuProPheGlyGluSerValMetAspLeuValIleGly 114
DB 394 GTGATAAGTTCCTCCATATCAATGTTGATGATGCTTGGATCATGCTGGTTATTGGT 453

QY 115 CysGlyProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGly 134
DB 454 TGTGCTCTCTGCTGTTAGCTTGGCTGCAGNATCAGCTAAGCTTGGATTAAAGTTGGA 513

QY 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLys 154
DB 514 CTCATTGTCAGATCTCTCTTTACTAACAAATACGGTGTGGGAAGATGAATTCAT 573

QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTyrLysAspThrIleValTyrLeuAsp 174
DB 574 GATCTGGGCTGCAAAATGATTGAGCATGTTTCGAGAGAGACTATTGTGTATCTGGAT 633

QY 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHis 194
DB 634 GATGACAGCCTATTACCATGGCGCTGCTTATGGAAGATTAGTCGAGCTTGTCTCCAT 693

QY 195 GluGluLeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGlu 214
DB 694 GAGGAGCTTTTGGAGGCTGTCGAGTCAGGTGCTCTGCTACCTTAGCTCGAAAGTTGAC 753

QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheLe 234
DB 754 AGCATTAACAGAGCTCTCATGCGCTTAGACTTGTGCTGTGACGACAATAACGTCATT 813

QY 235 ProCysArgLeuAlaThrValAlaSerGlyAlaLysSerGlyLysLeuLeuGluTyrGlu 254
DB 814 CCTGCGAGCTTGCCACGTGTGCTTCTGAGCAGCTTCGGGAAGCTCTTGCAATACGAA 873
  
```

```

QY 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluAsn 274
DB 874 GTTGGTGGACCTAGAGTCTGTGTGCAAACTGCATACGGCGGTGAGGTGAGGTGGAAT 933

QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLys 294
DB 934 AGTCCATATGATCCAGATCAAAATGGTTTTCATGGATTACAGAGATTATATAACGAGAA 993

QY 295 LeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThr 314
DB 994 GTTCGAGCTTAGAAGCTGAGTATCCACGTTTCTGTAGCCATGCTATGACAAGTCA 1053

QY 315 ArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
DB 1054 AGACTCTCTTCGAGAGACATGTTGGCTCAAAAGATGTCATGCCCTTTCATTGCTVA 1113

QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
DB 1114 AAACGAAAGCTCATGTTAAGATTAGATACACTCGGAATTCGAATTCGAAAGACTTACGAA 1173

QY 355 GluGluTrpSerTyrIleProValGlySerLeuProAsnThrGluGlnLysAsnLeu 374
DB 1174 GAGGAGTGGTCTATATCCACGTTGTTGTTGCTTCCCAACACCGCAACAAAGATCTC 1233

QY 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394
DB 1234 GCCTTTGGTGTGCTGCTAGCATCCGCAACAGCGCTATTCAGTTGTGAGATCT 1293

QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspLysSer 414
DB 1294 TTGCTGTAAGCTCCAAATATGTCATCAGTCATCGCAGAGATACATAAGAGAGAGACTACC 1353

QY 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
DB 1354 AAACAGATCAACAGT-----ATATTTCAAGCAACGCTTGGGTACT 1395

QY 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeu 454
DB 1396 TTATGCCACCAAGAAAGAAAGACAGAGCATTTCTTCTCTTGTGCTTGCACCTATA 1455

QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
DB 1456 GTTCAATTTCGATCGAAGGCAATAGAACGTTCTTCCGTACTTCTTCGCGCTTCCAAA 1515

QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
DB 1516 TGGATGTGCAAGGTTTCTAGGATCAACATTAACATCAGGAGATCTCGTTCTCTTGT 1575

QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
DB 1576 TTATACATGTTGCTCATTTCCACAAACAATTTGAGAAAGGTCATCATCATCTCATC 1635

QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
DB 1636 TCTGATCCAAACGGAGCAACCATGATATAAAACCTATCTCAAA 1677
  
```

RESULT 2

US-09-201-641-5
 ; Sequence 5, Application US/09201641A
 ; Patent No. 6232530
 ; GENERAL INFORMATION:
 ; APPLICANT: Cunningham Jr, Francis X
 ; APPLICANT: DellaPenna, Dean
 ; TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
 ; TITLE OF INVENTION: Mariogolds
 ; FILE REFERENCE: Quest 41-162
 ; CURRENT APPLICATION NUMBER: US/09/201,641A
 ; CURRENT FILING DATE: 1998-11-30
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1887

US-08-908-758-2

Alignment Scores:
 Pred. No.: 1,39e-131 Length: 960
 Score: 1202.00 Matches: 224
 Percent Similarity: 88.38% Conservative: 27
 Best Local Similarity: 78.87% Mismatches: 31
 Query Match: 43.82% Indels: 2
 DB: 4 Gaps: 1

US-09-701-395A-23 (1-529) x US-08-908-758-2 (1-960)

QY 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
 DB 66 AGTGTGATGATATCGACAAATTCGGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125
 QY 124 AlaGluAlaAlaLysLeuGlyValGlyLeuIleGlyProAspLeuProPheThr 143
 DB 126 GCAGATACACAGGTTAGTCTCAAGTTGGACTCATTGGTCTGATCTCTCTCTCTCTCT 185
 QY 144 AsnAsnTyrGlyValTyrGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
 DB 186 AACAACTACGGTGTGGGAAGATGAGTCAACGATCTTGCTGTCGCAAAATGATTGAG 245
 QY 164 HisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArg 183
 DB 246 CATGTTGGAGAGATACCTTGTGTATCTGAGCATGACAACTCTATTACCATTTGGTCT 305
 QY 184 AlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArgCysValGlu 203
 DB 306 GCTATGGAAGAGTGTAGTACCTTACTTACGAGAGGCTTGTGAGGAGGTGTGGAG 365
 QY 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
 DB 366 TCAGGTGTCTCGTATCTTACCTCCAAAGTTGAGAGCATACAGAGCTCTCTGATGGCTT 425
 QY 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
 DB 426 AGCTTTGTTTCCGTGTGAACAAAACACCTTGTTCCTGTCAGGCTTGTGCTGCTTCT 485
 QY 244 GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGln 263
 DB 486 GGAGCAGCTCTGGGAGCTTGTGCATACAAAGTTGGAGGCTTAGAGTCTGTGTCCAA 545
 QY 264 ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283
 DB 546 ACTGCTTACGGCTTGGAGTTGAGTGGTGAAGAGTCCATATGATCCAGAGCAGATGGTG 605
 QY 284 PheMetAspTyrArgAspTyrMetGlnLysLeuGlnCysSerGluGluGluTyrPro 303
 DB 606 TCCATGGATTACAGAGATTATACAAACGAGAAATCCGAGCTTAGAAGCTGAATATCCA 665
 QY 304 ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu 323
 DB 666 ACCTTTCTACCCCATGCTTATGACAAACAGCAGAGTCTCTTTGAGGAGACATGCTT 725
 QY 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
 DB 726 GCTTCAAAAGATGCTATGCTGCTTGAATTCCTTAAAGAAAGCTCTTGTGTGAGATTAG 785
 QY 344 ThrLeuGlyIleGlnValThrLysValTyrGluGluThrPheSerTyrIleProValGly 363
 DB 786 ACCTCGGAATCCGAATCTAAAGACTTACGAAGAGAAATGGTCTTATATATCCAGTAGGT 845
 QY 364 GlySerLeuProAsnThrGluGlnLysAspLeuAlaPheGlyAlaAlaAlaSerMetVal 383
 DB 846 GGTTCCTGCAACACGGAACAAAGAAATCTCGCTTTGGCTGTCAGCTAGCATGGTA 905
 QY 384 HisProAlaThr 387
 DB 906 CATCCCGCAACA 917

RESULT 5

US-08-399-561-5

; Sequence 5, Application US/08399561
 ; Patent No. 5792903
 ; GENERAL INFORMATION:
 ; APPLICANT: Hirschberg, Joseph
 ; APPLICANT: Cunningham Jr., Francis X.
 ; APPLICANT: Gantt, Elisabeth
 ; TITLE OF INVENTION: Lycopene Cyclase Gene
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kohn & Associates
 ; STREET: 30500 No. 5792903thwestern Highway, Suite 410
 ; CITY: Farmington Hills
 ; STATE: MI
 ; COUNTRY: US
 ; ZIP: 48334
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/399,561

; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kohn, Kenneth I.
 ; REGISTRATION NUMBER: 30,955
 ; REFERENCE/DOCKET NUMBER: P-305 (Hebrew Univ.)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 810-539-5050
 ; TELEFAX: 810-539-5055
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1650 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; ORGANISM: Lycopodium esculentum
 ; STRAIN: cv. VF36
 ; TISSUE TYPE: Leaf
 ; US-08-399-561-5

Alignment Scores:
 Pred. No.: 4.13e-78 Length: 1650
 Score: 750.50 Matches: 169
 Percent Similarity: 56.16% Conservative: 91
 Best Local Similarity: 36.50% Mismatches: 168
 Query Match: 27.36% Indels: 35
 DB: 1 Gaps: 11

US-09-701-395A-23 (1-529) x US-08-399-561-5 (1-1650)

QY 73 GlyGlySerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGlyGlnAla 92
 DB 265 GGTAGTAGTAGTGCCTTTTGTAGAGCTTGTACCTGACGACAAAGAGAGAGATCTGATTTT 324
 QY 93 LysLeuAlaAspLysLeuProPhe-----ProPheGlyGluSerValMetAspLeu 110
 DB 325 GAGCTT-----CCTATCTATGACCTTCAAAAGGGGTTCTTGTGGATCTT 369
 QY 111 ValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGly 130
 DB 370 GCTGTGGTGTGGTGGCCCTGCGAGACTTGTCTGTGTGACACCAAGCTTCTGACGAGGA 429
 QY 131 LeuLysValGlyLeuIleGlyProAsp-----LeuProPheThrAsnAsnTyrGlyVal 148
 DB 430 CTCCTCTGTTTGTTCATTTGATCCGAATCTCTAAATTTGATATGCTTAATAAATGCTGTT 489
 QY 149 TrpGluAspGluPheLysAspLeuGluArgCysIleGluHisAlaTrpLysAsp 168

Db 430 TGGGTGATGAATTTGAGGCTATGCTGTTAGATTGTCTAGATGCTACCTGCTGGT 549
Qy 169 ThrileValTyrLeuAspAsnAlaProValLeuIleGlyArgAlaTyrGlyArgVal 188
Db 550 GCACAGTGTACATTGATGATAATGCGCTAAAGATCTTCATAGACCTTATGGAAGGTT 609
Qy 189 SerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGluSerGlyValSer 208
Db 610 AACCGGAACAGCTGAATCGAAATGATGCGAAGATGATATGAATGCTGTTAAATTC 669
Qy 209 LeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValCys 228
Db 670 CACCAAGCCAAAGTTATAAGGTGATTCATGAG---GAATCGAATCCATGTTGATGTC 726
Qy 229 GluAsnGluPheIleProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGly 248
Db 727 AATGATGGTATTACTATTGAGCAGCGGTGGCTCGATGCAACATGCG---TTCCTAGA 783
Qy 249 LysLeuLeuGluTyrGluValGly---GlyProArgValCysValGluThrAlaTyrGly 267
Db 784 TCTCTTGTTCAGTATGATAAGCTTATAACCCCGG-----TATCAAGTTCCTATGTC 837
Qy 268 ValGluValGluValGluAsnProTyrAspProAsnLeuMetValPheMetAspTyr 287
Db 838 ATTTGGCTGAAGTGAAGACAGACCCCTTTCATGTAACAGATGGTTTCATGATGG 897
Qy 288 ArgAsp---TyrMetGlnLysLeuGlnCysSerGlu-----GluGluTyrProThr 304
Db 898 CGAGATTCTCATTTGAAGAACATATCTCAAGCAGAGAAATAGTAGAATACCAACT 957
Qy 305 PheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAla 324
Db 958 TTTCTTATGCAATGCCATTTTCACCAAGGATATTTCTTGAAGAAACATCCTCGTA 1017
Qy 325 SerLysAspAlaMetProPheAspLeuLysArgLysLeuMetSerArgLeuLysThr 344
Db 1018 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
Qy 345 LeuGlyIleGlnValThrLysValTyrGluGluGluTyrSerTyrIleProValGlyGly 364
Db 1078 TTGGGATAAAGTGAAGACATTTGAAGAGATGAATGCTATACCAATGGTGGT 1137
Qy 365 SerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyValAlaAlaSerMetValHis 384
Db 1138 CCATCTCCAGTATTACCTCAGAGAGCGTGTGAATCGGTGGTACAGCTGGCATGTTCT 1197
Qy 385 ProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerVal 404
Db 1198 CCATCACCAGTATATGTCGCAAGGACACTAGCTGCGGCTCCCTGTTGTCACATGCC 1257
Qy 405 IleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValSerGlyClnSerSerAla 424
Db 1258 ATAATTCATACCTCGGTTCTGAAGAAGT-----CATTCGGGT 1296
Qy 425 ValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArg 444
Db 1297 AATGAATATCACAGCTGTTGGAAGATTTGTGGCTATAGAGAGAGAGCTCAAGA 1356
Qy 445 AlaPheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThr 464
Db 1357 GAGTTCTTCTGCTCGGTATGATATCTCTGAGAGCTGATTTACCTGCTACAGAAGG 1416
Qy 465 PhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSer 484
Db 1417 TCTTTGATGATCTTTTGACTTAGAACCTCGTTATGGCATGCTCTTATGCTCGTA 1476
Qy 485 LeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSer 504
Db 1477 TTGTTTCTACCTGACTCATAGTTTTGGCTGCTCTCTATTTCTCTCATGCTCAATACT 1536
Qy 505 MetArg-----MetSerLeuValArgHisLeu 513
Db 1537 TCTAGATTTGAGATAATGACAAAGGAAGTGTTCATTAGTAATATGATCAACAATTTG 1596

Qy 514 LeuSerAsp 516
Db 1597 TTACAGGAT 1605
RESULT 6
US-09-201-641-1
; Sequence 1, Application US/09201641A
; Patent No. 6232530
; GENERAL INFORMATION:
; APPLICANT: Cunningham Jr, Francis X
; APPLICANT: Dellapenna, Dean
; TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
; TITLE OF INVENTION: Marigolds
; FILE REFERENCE: Quest 41-162
; CURRENT APPLICATION NUMBER: US/09/201,641A
; CURRENT FILING DATE: 1998-11-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Tagetes erecta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (304)..(1836)
; OTHER INFORMATION: beta-cyclase
US-09-201-641-1
Alignment Scores:
Pred. No.: 6,31e-78 Length: 1959
Score: 750.00 Matches: 180
Percent Similarity: 55.85% Conservative: 97
Best Local Similarity: 36.29% Mismatches: 189
Query Match: 27.34% Indels: 30
DB: 4 Gaps: 15
US-09-701-395A-23 (1-529) x US-09-201-641-1 (1-1959)
Qy 25 SerSerLysLeuAlaTyrAsnIleHisArgTyrGly-----SerSerCysArgValAsp 42
Db 352 AGTAACAAATTTGCGAAGAAATTTGAACAATTTGAATCAATTTGAATCAATCAATCAATCAAT 411
Qy 43 PheGlnValArgAlaAspGlyGlySerGlySerArgSerSerValAlaTyrLysGluGly 62
Db 412 TTTCAA-----GACTTTAGATTGTCGCCCAAAATATCC---CAATTCAAATTAGGG 459
Qy 63 PheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeu---PheValGln 81
Db 460 -----CAAAATATGCTGTTAAAGCTAGTAGTGTGCTTTGTTAGAACTTTGTTCT 510
Qy 82 MetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProIle 101
Db 511 GAAATCAAGAAAGAAATCTTGATTTGATCTTCCTATGATGAT----- 555
Qy 102 ProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeuSer 121
Db 556 CCATCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
Qy 122 LeuAlaAlaGluAlaLysLeuGlyLysValGlyLeuLeu-----GlyProAsp 139
Db 616 GTGGCTCAACAAGTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Qy 140 LeuProPheThrAsnAspTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGlu 159
Db 676 CTCATTGGCCCAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
Qy 160 ArgCysIleGluHisAlaTyrLysAspThrIleValTyrLeuAspAsnAlaProVal 179
Db 736 GATTTGTTGGATACAACTTGTGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Qy 180 LeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLys 199
Db 180 LeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLys 199

us-09-701-395a-23.rni

Thu May 22 09:39:17 2003

```

796 AGTCTTATAGACCATATGCAAGAGTCAATAGAAACAACTTAAACAAAGATGTTACAA 855
Db
200 ArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAla 219
QY
856 AAGTGATAGCAAAATGGGGTAAAGTTTCATCAAGCAAAAGTCAATCAAGTGAATTCATGAA 915
Db
220 GlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAla 239
QY
916 ---GAGTTAAATCTTTGTTGATTTGTTGATGATGATGATGATGATGATGATGATGATGAT 972
Db
240 ThrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGly---GlyPro 258
QY
973 CTGTGATCACTGGT---TTTCAAGATCTTTAGTTCAATATGATGAAGCCCTTATACCT 1029
Db
259 ArgValCysValGluThrAlaTyrGlyValGluValGluValGluValGluValGluVal 278
QY
1030 GGG-----TACCAAGTGGCTTATGGGATTTTACCGAAGTTTCAAGAACACCCCTTTTGAC 1083
Db
279 ProAsnLeuMetValPheMetAspTyrArgAsp---TyrMetGlnGlnLysLeuGln--- 296
QY
1084 GTTGATAAATGTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143
Db
297 ---CysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArg 315
QY
1144 AAAGCTAGAAATTCAGAAATCCCACTTTTATACCGGATGCCATTTTGTCTACAAAGA 1203
Db
316 LeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLys 335
QY
1204 ATCTTCTTGAGAAACATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263
Db
336 ArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValThrGluGlu 355
QY
1264 GAAAGAATGGCTTACAGGCTTAAGCAATTTGGGATTAAGTAAAGTAAAGTAAAGTAAAG 1323
Db
356 GluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAla 375
QY
1324 GAAGTGTGTATCCCGATGGCGGGCCCTACCAAGTCTCCCTCAAGGGTCTTGTGA 1383
Db
376 PheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeu 395
QY
1384 ATAGGTGGTACAGCAAGATGGTGATCCGTCACCAAGGATACATGTTGGCAAGACGCTPA 1443
Db
396 SerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAla 415
QY
1444 GCAGCGCCCGGATTTGTTGAAAGTCAATATCCGGTATCTTAAACGAAAGAACTATG 1503
Db
416 TyrVal---ValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
QY
1504 GTGGCCGAGCTACCGGAGATGATTTAGCAGCGGAATA-----TGGAGAGAA 1551
Db
435 LeuTrpProLysGluArgGlyArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIle 454
QY
1552 TTGTGGCTTATTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATATG 1611
Db
455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
QY
1612 TTGAAGCTGATTTGGAGAGTACTAGAGGTTCTTTGATGCGTTTTCGACTTGGAACT 1671
Db
475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
QY
1672 CGTTATTGGCATGGGTTTGTGCTGAGGTTGTTTCTACCGGAGTTAGTACGCTTTGGG 1731
Db
495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
QY
1732 CTATCGCTTTTCGGTCACTGTTCCAAATACCTTGTAGATTTGAATTTATG 1779
Db

```

RESULT 7
 US-08-908-758-4/c
 ; Sequence 4, Application US/08908758
 ; Patent No. 6429356
 ; GENERAL INFORMATION:
 ; APPLICANT: Calgene, Inc.
 ; TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALTY OIL

NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Calgene, Inc.
 STREET: 1920 Fifth Street
 CITY: Davis
 STATE: CA
 COUNTRY: USA
 ZIP: 95616
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
 COMPUTER: IBM PC
 OPERATING SYSTEM: Windows NT 4.0
 SOFTWARE: Microsoft Word For Windows 7.0a
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/908,758
 FILING DATE: 8-8-97
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/024,145
 FILING DATE: 8-9-96
 ATTORNEY/AGENT INFORMATION:
 NAME: Donna E. Scherer
 REGISTRATION NUMBER: 34,719
 NAME: Carl J. Schwedler
 REGISTRATION NUMBER: 36,924
 REFERENCE/DOCKET NUMBER: 122-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (916) 753-6313
 TELEFAX: (916) 753-1510
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1590 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 US-08-908-758-4

Alignment Scores: 8,72e-77 Length: 1590
 Pred. No.: 739.00 Matches: 164
 Score: 56.70% Conservative: 90
 Percent Similarity: 36.61% Mismatches: 144
 Best Local Similarity: 26.94% Indels: 50
 Query Match: 4 Gaps: 10
 Db:

US-09-701-395A-23 (1-529) x US-08-908-758-4 (1-1590)
 QY 107 ValMetAspLeuValIleGlyCysGlyProAlaGlyLeuSerLeuAlaGluAla 126
 Db 1304 GTTGTGATCTAGCTATCGTTGGCGGGCCCGCTGTTAGCGGTGCTCAGCAAGTC 1245
 QY 127 AlaLysLeuGlyLysValGlyLeuIle-----GlyProAspLeuProPheThrAsn 144
 Db 1244 TCGAAGCTGGACTCTCTGCTGCTCCATCATCTCTCCCAACACTCATTTGGCCCTAAC 1185
 QY 145 AsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGluHis 164
 Db 1184 AACTACGGATTTGGGTTTGGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTT 1125
 QY 165 AlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArgAla 184
 Db 1124 ACCTGTCGCGCGCTCGTCTACATCAACAGCGGCTCCGAAAGAGACCTGTGCCGCGCT 1065
 QY 185 TyrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArgCysValGluSer 204
 Db 1064 TACGGGAGAGTGAACCGCTAAACAGCTCAATCCAGATGCTTCCAGAGTGCATCACCACAC 1005
 QY 205 GlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSer 224
 Db 1004 GGTGTGTAGATTCATCAGCTGAAAGTCACTGAGTGTGTTCCAGAG---GAGGTAACTCC 948
 QY 225 LeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGly 244

Db 947 ACTGTTGCTCAGTGGTGTGAAGATTCAAGCTTCTGTTCTTCACGCTACTGCT 888
 QY 245 AlaAlaSerGlyLysLeuLeuGluTyrGluValGly---GlyProArgValCysValGln 263
 Db 887 ---TTTCAAGATGCTTGGTTCAGTATGATAGCCCTATAACCCCTGGG-----TATCAA 837
 QY 264 ThrAlaTyrGlyValGluValGluValGluValGluValGluValGluValGluVal 283
 Db 836 GTAGCTTATGATATCTCTGAGGTCGATGCTACCCGTTTGTGATGATAGATGGT 777
 QY 284 PheMetAspTyrArgAsp-----TyrMetGlnGlnLysLeuGlnCysSer 298
 Db 776 TTCATGCTGAGAGATGAAGCATCTTCACGCGTACCTGAGGTTAAGAGCGGAACAGC 717
 QY 299 GluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePhe 318
 Db 716 -----AAGATCCCTACCTTCTGTACCGGATCCGCTTCTTCCACAGCATCTTCT 663
 QY 319 GluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeu 338
 Db 662 GAAGAGAGCTCTGTGGCTAGCGGCTCTGAAGATGAAGATATCCAAAGAGAGATG 603
 QY 339 MetSerArgLeuLysThrLeuGluValGlnValThrLysValTyrGluGluTyrPhePhe 358
 Db 602 GTTCGAGGCTGAACATTTGGGATCAACGTGAAGCATGAGGATGAGAGACGAGCTGT 543
 QY 359 TyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheClyAla 378
 Db 542 GTGATCCCTATGAGGCTCTTACCGGCTTTCCTCAAGAGTGTGTGGCATTTGGCGGT 483
 QY 379 AlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAla 398
 Db 482 ACGCGGGATGCTTCATCTTCGATGTTTACATGTTGCTAGGCTCTTTCGAGCTGT 423
 QY 399 ProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValVal 418
 Db 422 CCGATAGTTGCAACGCTATAGTCGG-----AsnIleSerMetGlnAlaTyrPheSerLeu 435
 QY 419 SerGlyGlnSerSerAlaVal-----AsnIleSerMetGlnAlaTyrPheSerLeu 435
 Db 386 GGTGTAACAACAACGCTTGAAGAGATGAGCTCTCGCTGAGGTGTGGAGACTTG 327
 QY 436 TrpProLysGluArgLysAlaGlnArgAlaPhePheLeuPheGlyLeuGluLeuVal 455
 Db 326 TGGCTATTGAGCGCGGACAGAGAGTCTTCTGTTTGGGATGATATTGCTG 267
 QY 456 GlnLeuAspIleGluAlaThrArgThrPheArgThrPhePheArgLeuProThrTrp 475
 Db 266 AAGCTTATTTGATGCTACTAGAGGTTCTTTGATGCTTCTTGACCTGGAACCGCGT 207
 QY 476 MetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMet 495
 Db 206 TACTGCGATGTTCTTGTGCTGAGGCTGTTCTCCCGGACCTGCTTCTTCGGCGTG 147
 QY 496 TyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal-----510
 Db 146 TCGCTCTTCCATGCTCTTAATACCTCGAGGTTGGAGATCATGACGAAGGAAGTGT 87
 QY 510 -----510
 Db 86 CTTCTTGAATGATCAACAATTTGGTACAGATAGAGCGCAATTCACGACACTGG 27
 QY 511 ArgHisLeuLeuSerAspProSer 518
 Db 26 CGGCGGTACTAGTGGATCCGAGC 3

RESULT 8
 US-08-399-561-4
 ; Sequence 4, Application US/08399561
 ; Patent No. 5792903
 ; GENERAL INFORMATION:
 ; APPLICANT: Hirschberg, Joseph

APPLICANT: Cunningham Jr., Francis X.
 APPLICANT: Gantt, Elisabeth
 TITLE OF INVENTION: Lycopene Cyclase Gene
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kohn & Associates
 STREET: 30500 No. 5792903thwestern Highway, Suite 410
 CITY: Farmington Hills
 STATE: MI
 COUNTRY: US
 ZIP: 48334
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/399,561
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kohn, Kenneth I.
 REGISTRATION NUMBER: 30,955
 REFERENCE/DOCKET NUMBER: P-305 (Hebrew Univ.)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 810-539-5050
 TELEFAX: 810-539-5055
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1614 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 ORGANISM: Nicotiana tabacum
 STRAIN: cv. xanthi
 TISSUE TYPE: Leaf
 US-08-399-561-4
 Alignment Scores:
 Pred. No.: 1,54e-76 Length: 1614
 Score: 737.00 Matches: 174
 Percent Similarity: 54.70% Conservative: 111
 Best Local Similarity: 33.40% Mismatches: 180
 Query Match: 26.87% Indels: 56
 DB: 1 Gaps: 15
 US-09-701-395A-23 (1-529) x US-08-399-561-4 (1-1614)
 QY 14 ProValTyrThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHis 33
 Db 106 CCAGTTCATGATGTTCTGTAAA-----GCTAGCTCCTTAACTGTAAAGCCCCAT 159
 QY 34 ArgTyrGlySerSerCysArgValAspPheGlnValArgAlaAspGlyGlySerGlySer 53
 Db 160 AAGTTGCTGCTTAGG-----AAATTTGTGAAATTTGGGTAAAGGGTT 204
 QY 54 ArgSerSerValAlaTyrLysGluGlyPheValAspGluAspPheIleLysAlaGly 73
 Db 205 -----TGTTTAAAGCTAAG 219
 QY 74 GlySerGluLeuLeu-----PheValGlnMetGlnGlnThrLysSerMetGluLysGlnAla 92
 Db 220 AGTAGTGCCTTTTGGAGCTGTACCTGAGACCAAAAGGAAATCTTGATTTGAGCTT 279
 QY 93 LysLeuAlaAspLysLeuProPheGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 112
 Db 280 CCTATGATGAC-----CCTTCAAAAGGCTTGTGTAGATCTAGCTG 324
 QY 113 IleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLys 132
 Db 325 GTTGTGTGGTGGACCGCTGGACTTGCAGTTCACAGCAGGTTTCGGAGGCTGGACTATCG 384

Qy	487	SerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArg	506
Db	1432	CTTCTGAGCTTATATTTTCGGGCTGTCCCTTTCTCTCGCGCTTCAAATACTTCTTAGA	1491
Qy	507	-----MetSerLeuValArgHisLeuLeuSer	515
Db	1492	ATAGAGATTATGACAAAGGGAAGCTCTTCCTTTGGTAAATATGATCAACAAATTTGTTACAG	1551
Qy	516	Asp 516	
Db	1552	GAT 1554	
RESULT 9			
US-09-134-607A-11			
; Sequence 11, Application US/09134607A			
; Patent No. 6252141			
; GENERAL INFORMATION:			
; APPLICANT: Joseph Hirschberg et al.			
; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION			
; TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE			
; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID			
; TITLE OF INVENTION: BIOSYNTHESIS			
; NUMBER OF SEQUENCES: 25			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina			
; STREET: 20001 Jefferson Davis Highway, Suite 207			
; CITY: Arlington			
; STATE: Virginia			
; COUNTRY: United States of America			
; ZIP: 22202			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk			
; COMPUTER: Twinhead, Slimnote 890TX			
; OPERATING SYSTEM: MS DOS version 6.2,			
; OPERATING SYSTEM: Windows version 3.11			
; SOFTWARE: Word for Windows version 2.0,			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/134,607A			
; FILING DATE:			
; CLASSIFICATION: 800			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Friedman, Mark M.			
; REGISTRATION NUMBER: 33,883			
; REFERENCE/DOCKET NUMBER: 325/12			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 972-3-562553			
; TELEFAX: 972-3-562554			
; TELETYPE:			
; INFORMATION FOR SEQ ID NO: 11:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2897			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
US-09-134-607A-11			
Alignment Scores:			
Pred. No.: 2,08e-74 Length: 2897			
Score: 722.50 Matches: 165			
Percent Similarity: 58.94% Conservative: 79			
Best Local Similarity: 39.86% Mismatches: 151			
Query Match: 26.34% Indels: 19			
DB: 4 Gaps: 9			
US-09-701-395A-23 (1-529) x US-09-134-607A-11 (1-2897)			
Qy	104	GlyCysValMetAspLeuValValIleGlyCysGlyProAlaGlyLeuSerLeuAla	123
Db	1463	GGTCGGCGCTCAATTCGACGCTGATCATATATCGGAGCTGGCGCTTCGAGGTAGCT	1522

Dd 2995 CTTTTCGGACATGCTCAAAATTGCAGTGAGGTGGATAATTGTT 3036

RESULT l1
US-09-134-607A-8
; Sequence 8, Application US/09134607A
; Patent No. 6252141
; GENERAL INFORMATION:
; APPLICANT: Joseph Hirschberg et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION
; TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE
; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID
; TITLE OF INVENTION: BIOSYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 20001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead, Slimnote 890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; SOFTWARE: Word for Windows version 3.11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,607A
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedmam, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 325/12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-134-607A-8

Alignment Scores:
Pred. No.: 1.85e-74 Length: 1666
Score: 719,50 Matches: 164
Percent Similarity: 59.17% Conservatives: 78
Best Local Similarity: 40.10% Mismatches: 148
Query Match: 26.23% Indels: 19
DB: Gaps: 9

US-09-701-395A-23 (1-529) x US-09-134-607A-8 (1-1666)

Qy 109 AspLeuValIleGlyCysGlyProAlaGlyLeuSerLeuLaLaGLuLaLaLys 128S
||| :||||| ||||| ||||| ||||| |::|
Db 247 GACGTGATCATTAATCGGAGCTGGCCCTGCTGGGCTCAGGCTAGCTGAACAAGTTCTCAA 306E

Qy 129 LeucGlyLeuLyValGlyLeulleGlyProAspleuPhrothr -----AsnAsn 145P
||| :||||| ||||| ||||| ||||| |::|
Db 307 TATGGTATTAGGTATGTTGTGTGCCCTTCAC---CCACTCTCCAATGTGGCCAAAATAAT 363D

Qy 146 TyrGlyValTrpGlUAspGluPheLysAspleuGlyLeuGluArgCysilleGluHisAla 165H
||| :||||| ||||| ||||| ||||| ||||| |::|
Db 364 TAGCGTTTTGGGTGATGAGTTGAGAATTAGGACTGGAAAAATGCTTTAGATCATAA 423K

Qy 166 TrpLysAsphThrIleValTyrLeuAspAsnAspAlaProValLeulleGlyArgAlaTyr 185V
||| :||||| ||||| ||||| ||||| ||||| |::|

Alignment Scores:			
Pred. No.:	4,63e-74	Length:	2876
Score:	719.50	Matches:	164
Percent Similarity:	59.17%	Conservative:	78
Best Local Similarity:	40.10%	Mismatches:	148
Query Match:	26.23%	Indels:	19
DB:	4	Gaps:	9
US-09-701-395A-23 (1-529) x US-09-134-607A-14 (1-2876)			
QY	109	AspLeuValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLys	128
Db	1457	GAGCTGATCATTCGGCAGGTGGCCCTGCTGGCTCAGCTAGCTGAACAAGTTCTAAA	1516
QY	129	LeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr	145
Db	1517	TATGGTATTAGGTATGTTGTGTGGACCTTCA---CCACTCTCCATGTGCCCAATAAT	1573
QY	146	TyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArGlyCysIleGluHisAla	165
Db	1574	TATGGCTTTGGGTTGATCAGTTTGAGAAATTTAGGACTGGAAATTTGTTAGATCAATAA	1633
QY	166	TrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArGlaTyr	185
Db	1634	TGGCCTATGACTGTGTGCATATAATGATGACAACTAAGTATTGGGAAGACCATAT	1693
QY	186	GlyArgValSerArgHisLeuLeuHisGluGluLeuLysArgCysValGluSerGly	205
Db	1694	GGTAGACTTTAGTAAAGAAAGACTGAAGTTGAAATTTGTAATAGTTGTGTGGAGAACA	1753
QY	206	ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu	225
Db	1754	GTTGAAGTTTATAAGACTAAGTTTGGAAAGTG---GAACATGAAGAATTTGAGTCTTCA	1810
QY	226	ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla	245
Db	1811	ATTGTTTGTGATGATGTGAAGAAGATAAGAGGTAGTTTGGTTGGATGCAAGTGGTTTT	1870
QY	246	AlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArg---ValCysValGlnThr	264

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

```

COMPUTER: IEM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,598
FILING DATE:
CLASSIFICATION: 800
CLASSIFICATION: C12N 15/82
CLASSIFICATION: C12N 9/00
CLASSIFICATION: A01H 5/00
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9403943.5
FILING DATE: 01-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 94400626.1
FILING DATE: 23-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Calvetti, Frederick F.
REGISTRATION NUMBER: 28,557
REFERENCE/DOCKET NUMBER: GROFO 7000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 414-4000
TELEFAX: (202) 414-4040
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1756 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 67..1560
US-08-702-598-1

Alignment Scores:
Pred. No.: 8,94e-73
Score: 705.50
Percent Similarity: 57.95%
Best Local Similarity: 38.63%
Query Match: 25.72%
DB: 2

US-09-701-3995A-23 (1-529) x US-08-702-598-1 (1-1756)

```

QY	146	TyrGlyValTrpGluAspGluPheLysAspLeuGluArgCysIleGluHisAla	165
Db	430	TATGGCTGTTGGGTGAGTGGTAAAGTTCGGATTAGCAAGTGTCTAGATCATAAAG	489
QY	166	TrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArgGAlaTyr	185
Db	490	TGGCCCTGGAGTTGTGTTTCATATATAGTACACAAGACTAAAGTATTTGGACACACCATAT	549
QY	186	GlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGluSerGly	205
Db	550	GGTAGAGTAGTAGAAGAAGACTTGAAGTTGCAAAATGTGTGAATAGTGTGTGTGAANAATAGA	609
QY	206	ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu	225
Db	610	GTGAAGTGTATTAAGCCAAAGGTTTGAAGAGTGAAGCATGAA---GAAATTCAGTCTTCG	666
QY	226	ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla	245

Search completed: May 22, 2003, 00:17:48
Job time : 107 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 21, 2003, 22:12:08 ; Search time 359 Seconds
(without alignments)
3318.404 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRLNLISSCVWTFGT.....VRHLLSDPSGAVMVRAYLER 529

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O/cgn2_1/USPTO.spool/US0701395/runat_15052003_165749_3324/app_query.fasta_1.711
-DB-N_Geneseq_101002 -QFMT-fastap -SUFFIX-rng -MINMATCH=0.1 -LOOFCI=0
-LIST=45 -UNITS-bits -START=1 -END=-1 -MATRIX-blosum62 -TRANS-human40.cdi
-MODE=LOCAL -OUTFMT=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-USER=US0701395.ecgn_1.1.263.0runat_15052003_165749_3324 -NCPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2743	100.0	1898	21	AAZ45585	cDNA encoding an A
2	2728	99.5	1848	21	AAZ45595	cDNA encoding an A
3	2728	99.5	1848	21	AAZ46329	Adonis palaestina
4	1959.5	71.4	1860	18	AAZ45371	Arabidopsis thalia
5	1959.5	71.4	1860	18	AAZ45578	cDNA encoding an A
6	1934	70.5	1887	21	AAZ45584	Marigold epsilon-c
7	1931	70.4	1780	21	AAZ46328	Marigold lettuce ep
8	1892	69.0	2426	22	AAZ47465	Spinach lycopene c
9	1518.5	55.4	1377	21	AAZ45586	cDNA encoding a po
10	1394	50.8	1272	19	AAZ45124	Nucleotide sequenc
11	1394	50.8	1272	24	AAZ47356	Brassica napus eps
12	1202	43.8	960	19	AAZ45123	Nucleotide sequenc
13	1202	43.8	962	20	AAZ45064	Brassica napus lyc
14	1202	43.8	962	24	AAZ47355	Brassica napus eps
15	1187	43.3	1039	21	AAZ44887	Arabidopsis thalia
16	897	32.7	720	20	AAZ45065	Brassica napus lyc
17	761	27.7	1942	18	AAZ42979	Capsicum annuum ly
18	750.5	27.4	1650	17	AAZ40207	Tomato lycopene cy
19	750	27.3	1959	21	AAZ407582	Marigold beta-cycl
20	748	27.3	1986	21	AAZ451297	Arabidopsis thalia
21	739	26.9	1590	19	AAZ45125	Nucleotide sequenc
22	739	26.9	1590	20	AAZ45066	Brassica napus lyc
23	739	26.9	1590	24	AAZ47357	Brassica napus bet
24	737	26.9	1614	17	AAZ40206	Tobacco lycopene c
25	733	26.7	1988	21	AAZ45558	Arabidopsis thalia
26	722.5	26.3	2897	21	AAZ451519	L. pennellii lycop
27	722.5	26.3	3265	21	AAZ451521	L. pennellii lycop
28	719.5	26.2	1666	21	AAZ451516	L. esculentum lyco
29	719.5	26.2	1739	21	AAZ451518	L. pennellii lycop
30	719.5	26.2	2876	21	AAZ451517	L. esculentum lyco
31	705.5	25.7	1756	16	AAZ403851	Capsanthin-capsu
32	703.5	25.6	1740	21	AAZ451520	L. esculentum ogc
33	637.5	23.2	1731	21	AAZ434973	Soybean lycopene c
34	606	22.1	4928	16	AAZ494735	Lycopene cyclase g
35	606	22.1	4928	17	AAZ40205	Cyanobacterial lyc
36	562.5	20.5	1358	17	AAZ434972	Corn lycopene cyc
37	549.5	20.0	563	22	AAZ44247	Physcomitrella pat
38	325	11.8	853	21	AAZ434974	Wheat lycopene cyc
39	155.5	5.7	21091	21	AAZ451523	N. meningitidis pa
40	155.5	5.7	34980	21	AAZ451544	Neisseria meningit
41	155.5	5.7	34980	12	AAZ41607	Neisseria meningit
42	137.5	5.0	1234	12	AAZ413722	Lycopene cyclase -
43	137.5	5.0	1235	17	AAZ40795	Lycopene cyclase c
44	137.5	5.0	1235	18	AAZ451547	Erwinia herbicola
45	137.5	5.0	1893	18	AAZ45207	Squalene epoxidase

ALIGNMENTS

RESULT 1
AAZ45585
ID AAZ45585 standard; cDNA; 1898 BP.
XX
AC AAZ45585;
XX
DT 06-APR-2000 (first entry)
XX
DE cDNA encoding an Adonis palaestina lycopene epsilon-cyclase.

Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
lethal photooxidative damage; colouring; ss.

Adonis palaestina.

Key Location/Qualifiers

QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 Db 1553 CTTGGGCTTCACATATCATCTTTCGATCTGCTCTGTTTCCATGTACATGTTGGTTTG 1612
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 Db 1613 GCCCCAACAGCAGGAGTGTCACTTGTGAGACATTTGCTTTCAGATCCCTTCGGTGCA 1672
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 Db 1673 GTTATGGTAAAGAGCTTACCTCGAAAGG 1699

RESULT 2

AAZ45595
 ID AAZ45595 standard; cDNA; 1848 BP.

XX
 AC AAZ45595;

DT 06-APR-2000 (first entry)

XX
 DE

XX cDNA encoding an Adonis palaeatina lycopene epsilon-cyclase.
 KW lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring; ss.

XX
 OS Adonis palaeatina.

XX
 FH Key Location/Qualifiers

FT CDS 116..1705

FT /*tag= a

FT /product= "lycopene epsilon-cyclase"

FT /transl_except= (pos: 278..280, aa: Ser)

FT /transl_except= (pos: 743..745, aa: Asp)

FT /transl_except= (pos: 806..808, aa: Glu)

FT /transl_except= (pos: 1166..1168, aa: Val)

FT /transl_except= (pos: 1682..1684, aa: Arg)

XX
 PN W09963055-A1.

XX
 PD 09-DEC-1999.

XX
 PF 02-JUN-1999; 99WO-US12121.

XX
 PR 02-JUN-1998; 98US-0088724.

XX
 PR 02-JUN-1998; 98US-0088725.

XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.

XX
 PI Cunningham FX, Sun Z;

XX
 WPI; 2000-116366/10.

XX
 DR P-PSDB; AAY54303.

XX
 PT New carotenoid biosynthesis genes used to provide transgenic plants

XX
 PT with altered carotenoid content

XX
 PS Disclosure; Fig 23; 134pp; English.

XX
 CC The present sequence encodes a lycopene epsilon-cyclase protein.

XX
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The

XX
 CC specification also describes other enzymes involved in this pathway,

XX
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and

XX
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible

XX
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene

XX
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in

XX
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation

XX
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the

XX
 CC invention are used to provide transgenic plants with altered

XX
 CC carotenoid content. These plants may, for example, may be protected

XX
 CC from lethal photooxidative damage or have altered colouring.

SQ Sequence 1848 BP; 522 A; 334 C; 420 G; 572 T; 0 other;

Alignment Scores:

Pred No.: 2, 8e-280 Length: 1848

Score: 2728.00 Matches: 524

Percent Similarity: 100.00% Conservative: 5

Best Local Similarity: 99.05% Mismatches: 0

Query Match: 99.45% Indels: 0

DB: 21 Gaps: 0

US-09-701-395a-23 (1-529) x AAZ45595 (1-1848)

QY 1 MetGluLeuLeuGlyValArgAsnLeuLeuSerSerCysProValTrpThrPheGlyThr 20
 Db 116 ATGGAACCTACTGGTGTTCGCAACCTCATCTCTCTGCCCCTGCTGGACCTTTTGAACA 175
 QY 21 ArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArg 40
 Db 176 AGAAACCTTAGTAGTTCAAACTAGCTTATAACATACATCATGATATGGTTCTTCTGTAGA 235
 QY 41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerSerValAlaTyrLys 60
 Db 236 GTAGATTTTCAAGTCAGGCGCTGATGGTGAACGGGAGGTAGAACCTCTCTGCTTATAAA 295
 QY 61 GluGlyPheValAspGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheVal 80
 Db 296 GAGGTTTTTGTGGACGAGGAGATTTTATCAAAAGCTGGTGGTTCTGAGCTTTTGTGTC 355
 QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
 Db 356 CAATATGCGAGCAACCAAGCTCTATGGAGAAACAGGCGCAAGCTCGCGATTAAGTTGCCACCA 415
 QY 101 IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
 Db 416 ATACCTTTCCGGAGATCTGTGATGGACTGGTGTGTAAATAGGTGTGGACCTGCTGCTCT 475
 QY 121 SerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
 Db 476 TCACCTGGCTGCAGAAAGCTCTAAGCTAGGCTTGAAGTTGGGCTTATTTGCTGCTATCTT 535
 QY 141 PropheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArg 160
 Db 536 CCTTTTACAATTAATTTAGTGTGTGGGAAGACGAGTTTCAAAGATCTTGGACTTTGAACGT 595
 QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180
 Db 596 TGTATCGAGCATGCTTGGAAAGGACACCATCGTATATCTTTGACAAATGATGCTCTCTGCTT 655
 QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArg 200
 Db 656 ATTGCTCGTCATATGGACGAGTTAGCCGGCATTTGCTGCATGAAGAGTTGCTGAAAGG 715
 QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
 Db 716 TGTGTGAGTTCAGGTGTATCATATCTGAATCTTAAAGTGAAGGATCATCTGAAGCTGGT 775
 QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
 Db 776 GATGCCATAGTCTTGTAGTTGTGAAACGACATCTTATCCCTTGCAGGCTTGTCTACT 835
 QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValAlaGlyProArgVal 260
 Db 836 GTTGATCTGGAGCAGCTTCAGGGAACCTTTTGGAGTATGAAGTAGGTGGCCCTCGTGT 895
 QY 261 CysValGlnThrAlaTyrGlyValGluValGluAsnAsnProTyrAspProAsn 280
 Db 896 TGTGTCCAACTGCTTATGGTGTGAGGTTCAGGTGGAGAACCAATCCATACCATCCCCAAC 955
 QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300
 Db 956 TTAATGGTATTTATGGACTACAGACATATATGCAACAGAAATTTACAGTCTCGGAAGAA 1015
 QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePheGluGlu 320

Db	1016	GAATATCCACATTTCTGTATGTCATGCCATGTGCGCAACAAGACTTTTTTTTGAGGAA	1075
Qy	321	ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer	340
Db	1076	ACCTGTTTGGCCTCAAAAGATGCCATGCTTTCATCTACTGAAGAAACCTAAATGTCA	1135
Qy	341	ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGlnTrpSerTyrIle	360
Db	1136	CGATTCAAGACTCTGGGTATCCAGGTATCAAAAATTTATGAAGAGGAATGGCTTATATT	1195
Qy	361	ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAla	380
Db	1196	CGTGTGGGGTCTTTACCAACACACAGACAAAGAACCTAGCAITTGCTGCTCGACA	1255
Qy	381	SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGlnAlaProLys	400
Db	1256	ACCATGGTGCAATCCAGCAACAGGCTATCGTGTGTCAGATCACTATCAGAAGCTCCAAA	1315
Qy	401	TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly	420
Db	1316	TATGCTCTCTGTAATTGCAAGATTTTGAAGCAAGATACTCTGCATATGTGGTTCTCGA	1375
Qy	421	GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg	440
Db	1376	CAAAAGCAGTCAGTAACATTTCAATGCAAGCATGGACAGCTTTTGCCAAAGAGCGA	1435
Qy	441	LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu	460
Db	1436	AAAGCTCAAGACCATCTCTTCTTCGGGTAGAGCTTATGTGCAGCTAGATATTGAA	1495
Qy	461	AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe	480
Db	1496	GCACCAGAACGTTCTTTAGAACCTCTTCGCGTTCGCACCTTGGAGTGGTGGGGTTC	1555
Qy	481	LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu	500
Db	1556	CTTGGGTCTCTACTATCATCTTCGATCTGTATGTTTTCATGTGTATGTTTTCATGT	1615
Qy	501	AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla	520
Db	1616	GCCCCGAACACATGAGGATGTCACTTGTGAGACATTTGCTTCAGATCCCTTCGTGTGA	1675
Qy	521	ValMetValArgAlaTyrLeuGluArg	529
Db	1676	GTTATGTTTAAAGCTTACCTTCGAAGG	1702

RESULT 3	
AAZ46329	
ID	AAZ46329 standard; CDNA; 1848 BP.
XX	
AC	AAZ46329;
XX	
DT	07-MAR-2000 (first entry)
XX	
DE	Adonis palaestina epsilon lycopen cyclase cDNA.
XX	
KW	Lycopene; epsilon cyclase; bicyclic epsilon carotene; biosynthesis;
KW	carotenoid; epsilon ring; luteoxanthin; protein production;
KW	transgenic plant; altered synthesis; overexpression;
KW	nutritional value; pharmacology; colour; cancer; antisense;
XX	accumulation; substrate; beta-carotene; vitamin A; ds.
OS	Adonis palaestina.
XX	
FF	
Key	Location/Qualifiers
CDS	116..1705
FT	/*tag= a
FFT	/product= "Adonis palaestina epsilon lycopen cyclase"
FTT	
FTF	
FTT	
PN	WN9961399-A1.
PD	02-DEC-1999.

25-MAY-1999; 99WO-US10461.
26-MAY-1998; 98US-0084222.
(UYMA-) UNIV MARYLAND BALTIMORE.
Cunningham FX;
WPI; 2000-062667/05.
P-PSDE; AAY52593.
New eukaryotic epsilon-lycopene cyclase and related DNA, used to regulate carotenoid synthesis in plants and other hosts, e.g. for treatment of cancer
Disclosure; Fig 6A; 38pp; English.
This sequence represents cDNA encoding epsilon lycopene cyclase from *Adonis palaeatina*. This enzyme adds two epsilon rings to lycopene to form bicyclic epsilon-carotene. Carotenoids with two epsilon rings are found in significant amounts in relatively few plants (romaine lettuce produces an abundance of such a carotenoid, luteoxanthin), while carotenoids with two beta rings are ubiquitous and those with one beta and one epsilon ring are common. The invention relates to epsilon lycopene cyclases capable of forming bicyclic epsilon-carotene, and also to methods of transforming known carotenoids into novel or rare products. Epsilon lycopene cyclase nucleic acids may be used for recombinant protein production and for the generation of transgenic plants that have altered patterns of carotenoid synthesis. Such plants may accumulate new or rare carotenoids or overexpress known carotenoids resulting in altered nutritional value, pharmacology or colour. Carotenoids isolated from such transgenic plants are useful therapeutically, e.g., for treatment of cancer. Alternatively, epsilon lycopene cyclase antisense nucleic acids can be used to cause accumulation of the substrates of the enzyme, specifically beta-carotene or derived products such as vitamin A. The isolated eukaryotic enzyme allows synthesis of epsilon, epsilon-carotene (which can be isolated from natural sources only in small amounts) on a large scale, enabling bulk production of derived carotenoids.

Alignment Scores:		2.8e-280	Length:	1848
Pred. No.:	Score:	2728.00	Matches:	524
Percent Similarity:		100.00%	Conservative:	5
Best Local Similarity:		99.05%	Mismatches:	0
Query Match:		99.45%	Indels:	0
Db:		21	Gaps:	0

US-09-701-395A-23 (1-529) x AAZ46329 (1-1848)	
QY	1 MetGluLeuLeuGlyValArgAsnLeuLeuSerSerCysProValTrpThrPheGlyThr 40
Db	116 ARGGAACACTACTGGTGTTCGCAACCTCATCTCTCTTCGCCCTGCTGGACATTTTGGAA 175
QY	21 ArgAsnLeuSerSerSerLysLeuAlaTrpAsnIleHisArgTyrGlySerSerCysArg 40
Db	176 AGAAACACCTTAGTAGTTCAAACACTAGCTTATACATACATCATGATGCTCTCTCTGTAGA 235
QY	41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerSerValAlaTrpLys 60
Db	236 GTAGATTTTCAACTGAGGCGCTGATGTGGACGCGGAGTAGAACTCTGTGCTTATATA 295
QY	61 GluGlyPheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheVal 80
Db	296 GAGGGTTTTGTGCACGAGGAGGATTTATCAAACTGGTGGTTCTGAGCTTTGTTTGTCT 355
QY	81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
Db	356 CAARTGCACCAACAAAGTCTATGAGAAACAGGCCCAAGCTGCCGATAAAGTTGCCACCA 415

QY 101 IleProPheGlyCysLeuSerValMetAspLeuValValIleGlyCysGlyProAlaGlyLeu 120
 DB 416 ATACCTTTTCGGAGAAATCTGTGAGGACTTGGTTGTAATAGGTTGTGGACCTGCTGCTT 475
 QY 121 SerLeuAlaAlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
 DB 476 TCACCTGGCTGCAAGAGCTCTAGCTAGGCTTGAAGAGTTGGCCTTATTTGGCTGATCTT 535
 QY 141 ProPheThrAsnAsnTyrGlyValTTPGluAspGluPheLysAspLeuGlyLeuGluArg 160
 DB 536 CCTTTTACAAATATATATGTTGTGGAGAGGAGCTTCAAGATCTTGGACCTGACGT 595
 QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180
 DB 596 TGTATCGAGCTGCTTGGAGAGCACCATCGTATATCTTACATCTTGCATGCTGCTGCTT 655
 QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLysArg 200
 DB 656 ATGGTGTGTCATATGAGAGGAGTACCGGCAATTCATGATGAAGAGTGTGCAAAAGG 715
 QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
 DB 716 TGTGTGAGCTAGGTATCATATCTGAATCTTAAAGTGAAGAGTCACTGAAGCTGGT 775
 QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
 DB 776 GATGGCATAGCTTGTAGCTTTGTGAAGACGACATCTTTATCCCTTGCAGGCTTGCTACT 835
 QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgVal 260
 DB 836 GTTGCATCTGGAGCAGCTTCAGGGAACCTTTGGAGTATGAAGTAGTGCCCTCGTGT 895
 QY 261 CysValGlnThrAlaTyrGlyValGluValGluValGluAsnProTyrAspProAsn 280
 DB 896 TGTGTCCAACTGCTTATGTTGGTGGAGGTTGAGTGGAGACAAATCCATACGATCCCAAC 955
 QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300
 DB 956 TTAATGGTATTTATGACCTACAGACTATATGCAACAGAAATACAGTCTCGGAAGAA 1015
 QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu 320
 DB 1016 GAATATCCAACTTCTCTATGTCATGCCCATGTGCCCAACAGAGCTTTTTTTTGAGGAA 1075
 QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLysArgLysLeuMetSer 340
 DB 1076 ACCTGTTGGCCTCAAAAGATGCCATGCTTCGATCTACTGAGAGAGAACTAATGTCA 1135
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTrpSerTyrIle 360
 DB 1136 CGATTGAAGACTCTGGGTATCCAAAGTTACAAAAATTTATGAAGAGGAATGGTCTTATAT 1195
 QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAla 380
 DB 1196 CTTGTTGGGGGTTCTTTACCAACACAGAGCAAAAGAACTAGCATTTGCTGTCAGCA 1255
 QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
 DB 1256 AGCATGGTGCATCCAGCAAGCTATTGCTGTTGACGATCACTATCAGAGCTCCAAA 1315
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
 DB 1316 TATGCTTCTGTAATTTGCAAGATTTTGAAGCAAGATAAATCTCATATGTTGTTCTGGA 1375
 QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
 DB 1376 CAAGCAGTGCAGTAAACATTTCAATGCAAGCATGGAGCAGTCTTTGGCCAAAGAGCGCA 1435
 QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
 DB 1436 AAACGTCAAGAGCATCTCTTTTCGGGTTAGACCTTATGTCAGCTAGTATATGAA 1495
 QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480

DB 1496 GCACCAAGAGCTTCTTTAGAACCTTCTTCGCTTGCCAACTTGATGTGGTGGGTTTC 1555
 QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 DB 1556 CTTGGGCTTCACATCATCTTTTCATCTTGTATGTTTTCATGTACATGTTGTTTGTG 1615
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 DB 1616 GCCCGAACAGCATGAGGATGTCATCTGTGAGACATTTGCTTTTCAGATCCTTCTGTCGCA 1675
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 DB 1676 GTTATGTTAAAGCTTACCTCGAAAGG 1702

RESULT 4

AAT95371
 ID AAT95371 standard; cDNA; 1860 BP.

AC AAT95371;

DT 14-APR-1998 (first entry)

XX Arabidopsis thaliana epsilon cyclase cDNA.

XX Epsilon cyclase; E-cyclase; carotenoid; biosynthetic enzyme;
 KW pigment; vector; pATeps; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers
 FT CDS 109..1683

FT polyA_signal 1732..1736

FT /*tag= a

FT /*tag= b

XX WO9736998-A1.

XX 09-OCT-1997.

XX 28-JAN-1997; 97WO-US00540.

XX 29-MAR-1996; 96US-0624125.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Cunningham FX, Sun Z;

XX WPI; 1997-503091/46.

XX P-PSDB; AAW35346.

XX Eukaryotic carotenoid biosynthetic enzymes and related genes -
 useful to control ratio of various carotenoid(s) in host and for
 production of novel carotenoid pigments

XX Claim 4; Page 33-35; 89pp; English.

CC This cDNA sequence codes for Arabidopsis thaliana epsilon cyclase
 (see AAW35346), an enzyme responsible for the formation of epsilon
 end-groups in carotenoids. It was isolated by mass excision and
 colour complementation screening of an A. thaliana cDNA library.
 CC Claimed expression vector pATeps comprising the epsilon cyclase
 gene is deposited at ATCC 98005. Claimed isolated DNA sequences
 (see AAW95371-73 and AAT95389-91) encoding the claimed eukaryotic
 carotenoid biosynthetic enzymes epsilon cyclase, beta-carotene
 hydroxylase and isopentenyl pyrophosphate isomerase (see
 AAW35346-51) are used in methods for augmenting the accumulation of
 carotenoids and for the production of novel and rare carotenoids
 in host cells. Methods are also provided for controlling the ratio
 of various carotenoids in a host, and for screening for eukaryotic
 genes that encode enzymes of carotenoid biosynthesis and metabolism.

XX Sequence 1860 BP; 531 A; 327 C; 448 G; 554 T; 0 other;

Alignment Scores:

Pred. No.:	2,14e-198	Length:	1960
Score:	1959.50	Matches:	376
Percent Similarity:	81.65%	Conservative:	60
Best Local Similarity:	70.41%	Mismatches:	81
Query Match:	71.44%	Indels:	17
DB:	18	Gaps:	5

US-09-701-395A-23 (1-529). x AAT95371 (1-1860)

[illegible]

Db	994	GTTCGGAGCTTAGAAGCTGAGTATCAACGTTTCTGTACGCCATGCCCTATGACAAAGTCA	1053
Qy.	315	ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu	334
Db	1054	AGACTCTCTTCGAGGAGACATGTTTGGCGCTCAAAGATGTCATGCCCTTGTATTGCTCA	1113
Qy	335	LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu	354
Db	1114	AAACCAAGCTCATGTTAGATTAGATACACTCGGAATTCGAATTCAAAGACTTACGAA	1173
Qy	355	GluGluTrrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeu	374
Db	1174	GAGGAGTGTGCTCATATCCAGTTGGTGTCTTCGCCAACACCGCAAAAGAATCTC	1233
Qy	375	AlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer	394
Db	1234	GCGTTTGTGTGTCGCGCTAGCTGGTACATCCCGCAACAGGCTATTCACTGTGAGATCT	1293
Qy	395	LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer	414
Db	1294	TTGTCTGAGCTCCAAATATGTCATCAGTCATCGCAGAGATACTAGAGAGAGACTACC	1353
Qy	415	AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer	434
Db	1354	AMCAGATCAACAGT-----AATATTTCAAGACAAGCTTGGGATACT	1395
Qy	435	LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIle	454
Db	1396	TTATGCCACCACAAAGGAAAGACAGAGAGCATTTCTTCTTTGTGTCTGCACATCA	1455
Qy	455	ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr	474
Db	1456	GTTCATTCGATACCGAAGGCATTAGAAAGCTCTTCGCTACTTTCTTCGCCCTTCCAAA	1515
Qy	475	TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer	494
Db	1516	TGGATGTGCAGGGGTTTCTAGGATCAACATTAACTACAGAGATCTCGTTCTCTTTGCT	1575
Qy	495	MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu	514
Db	1576	TTATACATGTTCTGCTATTTCACCAACAATTTGAGAAAGGTCATCAATCATCTCATC	1635
Qy	515	SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu	528
Db	1636	TCTGATCCACCGGAGCAACCATGATAAAACCATCTCTCAAA	1677

1677

RESULT 5

AAZ45578
ID AAZ4

XX
AC

HC
XX
DE

DT 06-2A XX

DE
XX
cDNAKW
Lycos
KW
LycosKW
KW
KW

KW XX leth

OS
XX
Arab:

Key

ET
ET
CDS

XX
LE

XX
N₂
WO996

09-DEF

02-JU

ID AAZ45578 standard; cDNA; 1860 bp.
XX

AAZ45578;

XX
DT 06-APR-20

XX (1185 bp)
DE cDNA encoding of 1185 bp

XX encoding an Arabidopsis thaliana lycopene epsilon-cyclase.

Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
isopentenyl pyrophosphate isomerase

dimethylallyl pyrophosphate isomerase; beta-carotene hydroxylase;
lethal photooxidation³
transgenic plant; carotenoid content

XX photooxidative damage; colouring; ss.

Atalapha thaliana, XX

FH	key	Location/Qualifiers
FT	CDS	109 1593

```
ET 105.11083
ET /*tag= a
```

```

/PRODUCT= "lycopene epsilon-cyclase"

```

W09963055-A1.

09-DEC-1999.

F 02-JUN-1999; 99WO-US12121.

```

XX PR 02-JUN-1998;      98US-0088724.
XX PR 02-JUN-1998;      98US-0088725.
XX PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX PI Cunningham FX, Sun Z;
XX DR WPI; 2000-116366/10.
XX DR P-PSDB; AAY54283.
XX CC
XX PT New carotenoid biosynthesis genes used to provide transgenic plants
XX PT with altered carotenoid content -
XX PS Example; Page 61-63; 134pp; English.
XX CC The present sequence encodes a lycopene epsilon-cyclase protein.
XX CC The enzyme is involved in the carotenoid biosynthesis pathway. The
XX CC specification also describes other enzymes involved in this pathway.
XX CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase,
XX CC beta-carotene hydroxylase, IPP isomerase catalyses the reversible
XX CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
XX CC hydroxylase is responsible for hydroxylating the beta-endgroup in
XX CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
XX CC of epsilon-endgroups in carotenoids. The polynucleotides of the
XX CC invention are used to provide transgenic plants with altered
XX CC carotenoid content. These plants may, for example, may be protected
XX CC from lethal photooxidative damage or have altered colouring.
XX SQ Sequence 1860 BP; 531 A; 327 C; 448 G; 554 T; 0 other;

Alignment Scores:
Pred. No.:          2,14e-198           Length:         1860
Score:             1959.50              Matches:         376
Percent Similarity: 81.65%               Conservative:    60
Best Local Similarity: 70.41%            Mismatches:     81
Query Match:       71.44%                Indels:        17
DB:                21                    Gaps:          5

US-09-701-395A-23 (1-529) x AAZ45578. (1-1860)

Qy   1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTtp 16
    ||||| :|||:|||||
Db 109 ATGGAGTCGTGTGGCGCTAGGAATTCGCACCAATGGCGGTTTCAACATTCCGTCATGG 168

Qy 17 ThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyArgSnileHisArgTyRgly 36
    :: :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 169 AGTTGTGCAAGAAATAATTTCCAGTGTTTAAGAGATACAGCTATAGGAATATTCGTTCCGCT 228

Qy 37 SerSerCysArgValaspPheGlnValArgAlaaspGlyGly-----SerglySerArg 54
    ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 229 ---TTGTGT-----AGTGTACAGCTACGCCGCGCGGAAGTTCCCGGTAGTGAG 273

Qy 55 SerSerValAlaTyRLysGluGlyPheValaspGluGluaspPheIleLysAlaGlyGly 74
    ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 274 AGTTGTGTAGCGGTGAGAGAAGATTTCCTGCACGAAGAAGATTTCGTGNAACCTGGTGTG 333

Qy 75 SerGluLeuLeuPheValGlnMetGlnInThrLysSerMetGluLysGlnAlaLysLeu 94
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 334 TCTGAGATCTATTGTTGTTCAAATGAGCAGACAAGAAGATATGGATGAACAGTCTAACGCTT 393

Qy 95 AlaaspLysLeuProIlePropheGlyGluSerValMetaspLeuValIleGly 114
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 394 GTTGATAGTTGTCCTTATCAATTGGTAGTGCTTTGGATGCATGTGTTATTGCT 453

Qy 115 CysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaAlaLysLeuGlyLeuLysValGly 134
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 454 TGtGStCtGtGtTgTtAgCcTtGgCtCAGATcAGTcAGCTtGGATtAAAGtTGA 513

Qy 135 LeuileGlyProaspLeuproPheThrAsnAsnTyRGlyValTrpGluaspGluPhetys 154
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 514 CTCATtGtGtCAGAtCTtCtTtTACTAcAAATtACGgtTtTGGAGAtGAATtCAAT 573

```

[illegible]

QY 397 GluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyr 416
 DB 1290 GAAGCTCCTTATATGACGACGATTAATGCAAGATTTAGGGAAGGAATTCAAAACAG 1349
 QY 417 ValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerLeuTrp 436
 DB 1350 ATGCTTGATCATGGAAGATACACCAACATCTCAAGCAAGCTTGGGAACACTTTGG 1409
 QY 437 ProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGln 456
 DB 1410 CCCTTTGAAGGAAGAACAGAGAGATCTTCTCTTTGGATAGCACTGATGTCACG 1469
 QY 457 LeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMet 476
 DB 1470 ATGGATATTAGGGGACCCGACATCTTCCGGACATCTTCCGCTGCCACATGGATG 1529
 QY 477 TrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyr 496
 DB 1530 TGGTGGGGTCTTCTGGATCTTCTGTTATCATCACTCACTGATGATAATATTGGCTTTAC 1589
 QY 497 MetPheValLeuAlaProAsnSerMetArgMetSerLeuValAlaArgHisLeuLeuSerAsp 516
 DB 1590 ATGTTTATCATAGCACCAGCATCCCTGAGAAATGGCTGTTAGACATTTGCTTCRGAC 1649
 QY 517 ProSerGlyAlaValMetValArgAlaTyrLeu 527
 DB 1650 CCGACAGGAGGAACAATGTAAAGCGTATCTC 1682

RESULT 7:

AZ46328
 ID AA246328 standard; cdna; 1780 BP.

XX AC AA246328;

XX DT 07-MAR-2000 (first entry)

XX DE Romaine lettuce epsilon lycopene cyclase cDNA.

KW Lycopene; epsilon cyclase; bicyclic epsilon carotene; biosynthesis;
 KW carotenoid; epsilon ring; lactucaxanthin; protein production;
 KW transgenic plant; altered synthesis; overexpression;
 KW nutritional value; pharmacology; colour; cancer; antisense;
 KW accumulation; substrate; beta-carotene; vitamin A; ds.

XX OS Lactuca sativa.

XX FH Key Location/Qualifiers
 XX FT 77..1677

XX FT /*tag= a
 XX FT /product= "Romaine lettuce epsilon lycopene cyclase"
 XX FT /transl_except= (pos:182..187, aa:Arg)
 XX FT /note= "This codon has an apparent 3 base insertion,
 XX FT leading to an insertion of an additional Arg
 XX FT residue in the protein"

XX FT 1457..1459
 XX FT /*tag= c
 XX FT /note= "Encodes Met"

XX PN WO9561399-A1.

XX PD 02-DEC-1999.

XX XX 25-MAY-1999; 99WO-US10461.

XX XX 26-MAY-1998; 98US-0084222.

XX XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX XX Cunnigham FX;

XX XX WPI; 2000-062667/05.

DR P-PSDB; AAY52592.

XX PT
 XX PT
 XX PS

New eukaryotic epsilon-lycopene cyclase and related DNA, used to regulate carotenoid synthesis in plants and other hosts, e.g. for treatment of cancer

Claim 4; Fig 3; 38pp; English.

This sequence represents cDNA encoding epsilon lycopene cyclase from romaine lettuce. This enzyme adds two epsilon rings to lycopene to form bicyclic epsilon-carotene. Carotenoids with two epsilon rings are found in significant amounts in relatively few plants (romaine lettuce produces an abundance of such a carotenoid, lactucaxanthin), while carotenoids with two beta rings are ubiquitous and those with one beta and one epsilon ring are common. The invention relates to epsilon lycopene cyclases capable of forming bicyclic epsilon-carotene, and also to methods of transforming known carotenoids into novel or rare products. Epsilon lycopene cyclase nucleic acids may be used for recombinant protein production and for the generation of transgenic plants that have altered patterns of carotenoid synthesis. Such plants may accumulate new or rare carotenoids or overexpress known carotenoids, resulting in altered nutritional value, pharmacology or colour. Carotenoids isolated from such transgenic plants are useful therapeutically, e.g., for treatment of cancer. Alternatively, epsilon lycopene cyclase antisense nucleic acids can be used to cause accumulation of the substrates of the enzyme, specifically beta-carotene or derived products such as vitamin A. The isolated eukaryotic enzyme allows synthesis of epsilon, epsilon-carotene (which can be isolated from natural sources only in small amounts) on a large scale, enabling bulk production of derived carotenoids.

SQ Sequence 1780 BP; 526 A; 335 C; 382 G; 536 T; 1 other;

Alignment Scores:

Pred. No.: 2 18e-195 Length: 1780
 Score: 1931.00 Matches: 375
 Percent Similarity: 78.51% Conservative: 56
 Best Local Similarity: 68.31% Mismatches: 78
 Query Match: 70.40% Indels: 40
 DB: 21 Gaps: 5

US-09-701-395a-23 (1-529) x AA246328 (1-1780)

QY 1 MetGluLeuGlyValArgAsnLeu-----IleSerSerCysProVal 15
 DB 77 ATGGAGTGTCTTGGAGCTCGAAACATGACGGCAACAATGGCGTTTATACGTGCCTAGA 136
 QY 16 TrpThr-----PheGlyThr 20
 DB 137 TTCACGGACTGTAATATACGGCACAAATTTTCGTACTGTGAACACGAAGATTTACTAAT 196
 QY 21 ArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArg 40
 DB 197 TTATCAGCATCGTCTTCTGCTCGCTCAAAATTAAGTCACGCGCTAAAGCGACCGTGTGTA 256
 QY 41 ValAspPheGlnValArgAlaAspGlySerGlySerArgSerSerValAlaTyrLys 60
 DB 257 GTGGAT-----AAA 265
 QY 61 GluGlyPhe-----ValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeu 78
 DB 266 CAAGGGATTTCCGTAGCAGACGAAGATTTATGGAAGCGCGGTGGATCGGAGCTGTTT 325
 QY 79 PheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeu 98
 DB 326 TTGTGTTCAATGACGAGGACTAAGTCCATGGAAGCCAGCTTAACCTTCGAAAGGTA 385
 QY 99 ProProIleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAla 118
 DB 386 GCACAGATACCAATTGGAAATTTGCATCTGCTGGTTGTAATCGGTGTGGCCCTGCT 445
 QY 119 GlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyPro 138
 DB 446 GGCCTTGTCTTGTGTCAGAGTCAGCCAAACTAGGCTTGAACGTTGGACTCATTTGGCCCT 505

-18 - CCGACGTTGTAGTAGAATGTAGTCTAGTCGTAGGAG

-----GCTTTTGAGGAGTGGC 377

```
QY 57 -----ValAlaTyrLysGluGlyPheValAspGluGluAspPheIle 70
Db 378 AATAGGCACAGTAGTGGCGTAATTCGCCACAGGATTTTCGGAACGAAGAAGATTTTCATC 437
QY 71 LysAlaGlyCysSerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLys 90
Db 438 AAGCTGGGTGGTTCGCCAGCTCTTTATGTTCAATGCAGCAATAAAGCTATGGATGTG 497
QY 91 GlnAlaLysLeuAlaAspLysLeuProProIleProPheGlyGluSerValMetAspLeu 110
Db 498 TACTCCAAAATTTCCGATAGCTGCGTCAAAATATCAGATGCCAATGAACCTGCTGGATAG 557
QY 111 ValValIleCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGly 130
Db 558 GTGGTTATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 617
QY 131 LeuLysValGlyLeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGlu 150
Db 618 TTTAAAGTTGGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 677
QY 151 AspGluPheLysAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIle 170
Db 678 GATGAATTCAGACCATTTGGAGCTGGAGGCTGTATCGAGCACCTTTGGCGTGATACCAT 737
QY 171 ValTyrLeuAspAsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArg 190
Db 738 GTGTATATTGATGATGACATCTATATATATATATATATATATATATATATATATATAT 797
QY 191 HisLeuLeuHisGluLeuLysArgCysValGluSerGlyValSerTyrLeuAsp 210
Db 798 CAATTACTTCACNAGGACTGGTGACAGGTGTTGGAGTCAGGTGCTCTTATCTGTAAT 857
QY 211 SerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsn 230
Db 858 GCGAAAGTGGAAATATTTATGGAAGACCTGATGGACATAGGCTGTGCTTGTGAACGT 917
QY 231 GluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaLysGlyLeu 250
Db 918 GGTGTCACTATTCCTCCGAGGCTGTAACCTGTTGCATCTGGAGCAGCTTCAGGGAAC 977
QY 251 LeuGluTyrGluValGlyProArgValCysValGlnThrAlaTyrGlyValGluVal 270
Db 978 CTGGAGTATCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1037
QY 271 GluValGluAsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyr 290
Db 1038 GAGGTGGAAACACAGCTTATGATGCCAATGTGATGCTGTTCATGGACTACAGAGACTAC 1097
QY 291 MetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetPro 310
Db 1098 ACTAAACTGAGCGTTCAATCTCTGGAGGCAAGTATCCAACTCTGTATGCAATGCCG 1157
QY 311 MetSerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetPro 330
Db 1158 ATATCAACCACTAGGATCTCTTTGAGGAGACCTGCTGGCTTCAGTAGATGCAATGCC 1217
QY 331 PheAspLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThr 350
Db 1218 TTTGACCTGCTCAAGAAAAAGCTTATGACAAGATTACAACTATGGGTGCTGTATCACC 1277
QY 351 LysValTyrGluGluGluTrpSerTyrIleProValGlySerLeuProAsnThrGlu 370
Db 1278 AAAATATATGAGAGAGAGTGGTCTTATATACCTGTTGGTGGTGGTGGTGGTGGTGGTGG 1337
QY 371 GlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSer 390
Db 1338 CAAGAAGAACTTGCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1397
QY 391 ValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLys 410
Db 1398 GTCGTGAGATCACTGTCAGAGCTCCAAAGATGCTCTGCAATTCGCAACTGATCAAG 1457
QY 411 GlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGln 430
```

```
Db 1458 AATGACCTGTCAAAAAATGCAATATTGCTCAGAGGAGTGTGGGAATATCTCAATGCAA 1517
QY 431 AlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGly 450
Db 1518 GCCTGGAAATACCTTTGGCCACAAAGAAAGACGTCAAGAGCATTCTTCTGTTCGGA 1577
QY 451 LeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePhe 470
Db 1578 CTATCACTTATAGTCCAGCTTGATATTGAGGTATCAGSACATTTCTCCGACCTCTTC 1637
QY 471 ArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeu 490
Db 1638 CGAGTGCACAAATGATGTGGGAGGATTCCTCGGTTCTTAATCTCTCTAGCTGATCTC 1697
QY 491 ValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
Db 1698 ATATTGTTGGCTTTTATATGTTCTTATGCTCCGAATGACTTGAGATGGGTCTTATA 1757
QY 511 ArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeu 527
Db 1758 AGGCATCTACTATCTGATCTACAGGGCGCACCATGATAAGAACGTACATA 1808
RESULT 9
AAZ45586
ID AAZ45586 standard; cDNA; 1377 BP.
XX
AC AAZ45586;
XX
DT 06-APR-2000 (first entry)
XX
DE cDNA encoding a potato lycopene epsilon-cyclase protein.
XX
KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
XX lethal photooxidative damage; colouring; ss.
XX
OS Solanum tuberosum.
XX
FH Key Location/Qualifiers
FT CDS 14..1150
FT /*tag= a
FT /product= "lycopene epsilon-cyclase"
XX
XX WO963055-A1.
XX
XX 09-DEC-1999.
XX
XX 02-JUN-1999; 99WO-US12121.
XX
XX 02-JUN-1998; 98US-0088724.
XX
XX 02-JUN-1998; 98US-0088725.
XX
XX (JYMA-) UNIV MARYLAND BALTIMORE.
XX
XX Cunnigham Fx, Sun Z;
XX
XX WPI; 2000-116366/10.
XX
XX P-PSDB; AAY54291.
XX
XX New carotenoid biosynthesis genes used to provide transgenic plants
XX with altered carotenoid content
XX
XX Disclosure; Fig 15A; 134pp; English.
XX
CC The present sequence encodes a lycopene epsilon-cyclase protein.
CC The enzyme is involved in the carotenoid biosynthesis pathway.
CC specification also describes other enzymes involved in this pathway,
CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
CC hydroxylase is responsible for hydroxylating the beta-endgroup in
```


CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.

XX
 SQ Sequence 1377 BP; 386 A; 255 C; 297 G; 431 T; 8 other;

Alignment Scores:
 Pred. No.: 1.32e-151 Length: 1377
 Score: 1518.50 Matches: 281
 Percent Similarity: 87.27% Conservativity: 48
 Best Local Similarity: 74.54% Mismatches: 47
 Query Match: 55.36% Indels: 1
 DB: 21 Gaps: 1

US-09-701-395A-23 (1-529) x AA245586 (1-1377)

Qy 151 AspGluPheLysAspLeuGluArgCysLeuGluHisAlaTrpLysAspThrIle 170
 Db 14 GATGAGTTCAAGATCTTGGTCTTCAAGCCTGCATTGAACATGTTTGGCGGATACCAATT 73
 Qy 171 ValTyrLeuAspAsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArg 190
 Db 74 GTATATCTTGATGATGATGATGATCTTATCTTATGCGCGCTATGGAAGAGTTAGTCGC 133
 Qy 191 HisLeuLeuHisGluGluLeuLysArgCysValGluSerGlyValSerTyrLeuAsp 210
 Db 134 CATTTACTGCACGAGGAGTTACTCAAAAGGTGTGTGGAGGAGGTTTGTATCTAATAC 193
 Qy 211 SerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsn 230
 Db 194 TCGAAGTGGATAGATTGTTGAGGCCCAAAATGCCACAGCTTGTAGAGTGGCAGGGT 253
 Qy 231 GluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeu 250
 Db 254 GATGTTGTGATCTCCTGCAGGTTTGTGACTGTTCGCGAGCAGCTCCGGGAAATC 313
 Qy 251 LeuGluTyrGluValGlyGlyProArgValCysValGluThrAlaTyrGlyValGluVal 270
 Db 314 TTGCAGTATGATGTTGGGAGGTCCTAGAGTTCTGTTCACACAGCTTATGGAGTGAAGTT 373
 Qy 271 GluValGluAsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyr 290
 Db 374 GAGTCTGATACCAATCCATTGTGACCCGAGCCTGATGTTTTCATGGATTATAGACTAT 433
 Qy 291 MetGlnGlnLysLeuGlnCysSerGluGluTyrProThrPheLeuTyrValMetPro 310
 Db 434 GTCAGACACGCGCTCAATCTTTAGAGCTAAATATCCACATTTCTATGCCATGCC 493
 Qy 311 MetSerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetPro 330
 Db 494 ATGCTCCACACGAGTCTTTTCGAGGAAATGTTTGGCTTCAAAAGATGCAATGCCA 553
 Qy 331 PheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThr 350
 Db 554 TTCGATCTGTTTAAAGAAATATGATGTTAGATTGACACCCCTCGGTGTAAGAATAAA 613
 Qy 351 LysValTyrGluGluGluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGlu 370
 Db 614 GAAATTTATGAGGAGGATGCTTACATACACAGTTGGAGATCTTGGCAATACAGAA 673
 Qy 371 GlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSer 390
 Db 674 CAAAAAACACTTGCATTTGGTGGCTGCTAGCTAGCTATCCAGCCACACAGGTTATTCA 733
 Qy 391 ValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLys 410
 Db 734 GTCGTGATGATCCTGCTGAGAGCTCCAAAATGCGCCTTCGTGTCGCAATATATATACGA 793
 Qy 411 GlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAlaSerMetGln 430
 Db 794 CAAATCATACCAAGATATGCTTACT--AGTTCAGTACCCCGAGTATTTCAACTCAA 850

Qy 431 AlaTrpSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGly 450
 Db 851 GCTTGGACACTCTTGGCCACAGAACGAAACACAAAGATCGTTTTTCTTATTTGA 910
 Qy 451 LeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePhe 470
 Db 911 CTGGCTCTGATATTGACGTGATATTGAGGGGATAAGTCAATTTTCCGGCGGTTCTTC 970
 Qy 471 ArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerLeuSerSerPheAspLeu 490
 Db 971 CGTGTGCCAAAATGGATGTGCAGGATTTCTTGGTTCAAGTCTTCTTNAGCAGACCTC 1030
 Qy 491 ValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
 Db 1031 ATGTTATTTGCTTCTACATGTTATTATTGCACCAATGCATGAGAGAGGCTTATC 1090
 Qy 511 ArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeu 527
 Db 1091 AGACATCTTTTATCTGATCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141

RESULT 10

AAV19124 standard; cDNA; 1272 BP.

AAV19124;

10-AUG-1998 (first entry)

Nucleotide sequence of epsilon cyclase cDNA clone 7-6.

Epsilon cyclase; SSU/crtB fusion protein; carotenoid; lycopene; lutein;
 zeaxanthin; canthaxanthin; vitamin; colourant; ss.

Brassica napus.

WO9806862-A1.

19-FEB-1998.

08-AUG-1997; 97WO-US14035.

09-AUG-1996; 96US-0024145.

(CALJ) CALGENE INC.

Shewmaker CK;

WPI; 1998-159551/14.

Use of constructs comprising a carotenoid biosynthesis gene - for
 producing plants and seeds having altered carotenoid levels,
 modified fatty acid compositions or altered tocopherol levels.

Example 1B; Fig 10; 70pp; English.

This is the nucleotide sequence of the Brassica napus epsilon cyclase
 cDNA clone 7-6. It was used in the method of the invention with
 the SSU/crtB fusion protein (AAV19122), to create plants and seeds with
 altered carotenoid levels. They can be used for the production of
 particular carotenoids, e.g. alpha-carotene, beta-carotene, lycopene,
 lutein, zeaxanthin, canthaxanthin, which can be used as supplements,
 particularly vitamin supplements, as vegetable oil based food products
 and food ingredients, as feed additives in animal feeds, as colourants
 or therapeutic agents.

Sequence 1272 BP; 342 A; 271 C; 311 G; 348 T; 0 other;

Alignment Scores:

Pred. No.:	2.15e-138	Length:	1272
Score:	1394.00	Matches:	265
Percent Similarity:	85.51%	Conservativity:	30
Best Local Similarity:	76.81%	Mismatches:	46

Query Match:	50.82%	Indels:	4
DB:	19	Gaps:	2
US-09-701-395A-23 (1-529) x AAV19124 (1-1272)			
Qy 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123			
Db 67 AGTGTGATGGATATCTGCAGANTCGGCTGTTGTGTGCTGCTGCTGGTTAGCTTGGCG 126			
Qy 124 AlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr 143			
Db 127 GCTGAATCAGCTAAGTTAGACTTAAAGTTGACATGATTGGTCTCTGACACTTCTTCACT 186			
Qy 144 AsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163			
Db 187 AACAACTACGGCTTTGGGAAGATGAGTTCACAGCATCTGGCTTGCCTGCAAAATGATTTCAG 246			
Qy 164 HisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArg 183			
Db 247 CATGTTGGAGAGATACCTTGTGTATCTGGACGATGACATCTTATACCATTTGGTTCGT 306			
Qy 184 AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu 203			
Db 307 GCTTATGGAAGAGTTAGTTCGACTTACTTACAGGAGGAGTTCTTCTGAGGAGGTGTGGAG 366			
Qy 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223			
Db 367 TCAGGTGTCGCTATCTTACGTCCTCAAGTTGAGACATCAACAGAAGCCTCTGATGGCCTT 426			
Qy 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243			
Db 427 AGCTGTGTTCTGTGAGCAAAACACCCCTGTGTCGTCGAGGCTTCCACATGTTCTTCT 486			
Qy 244 GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGln 263			
Db 487 GGAGCAGCTTCTGGGAAGCTCTTGCAATACGAAGTTGGAGGCCCTAGAGTCTGTGTCAA 546			
Qy 264 ThrAlaTyrGlyValGlnValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283			
Db 547 ACTGCTTACGGCTTGGAGGTGTGAGTGGAAAAGATGTCATATGATCCAGAGACATGGTG 606			
Qy 284 PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrPro 303			
Db 607 TTCATGATTTACAGATTTATACAAACGAGAAATCCGGAGCTTAGAGCTGAATATCCA 666			
Qy 304 ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu 323			
Db 667 ACGTTCCTCTACGCCATGCCTATGACAAAGACACAGAGCTCTTCTTTGAGGAGACATGCTT 726			
Qy 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343			
Db 727 GCTTCAAAAGATGTCATGCCCTTGATTGCTTTAAAAAGAGCTCTTCTTGCAGATTAGAG 786			
Qy 344 ThrLeuGlyIleGlnValThrLysValTyrGluGluGluTrpSerTyrIleProValGly 363			
Db 787 ACATCGGAATCGAATACTAAAGACTTACGAAGAGGAATGGTCTTATATCCAGTAGGT 846			
Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetVal 383			
Db 847 GGTTCCTTGCAAACACGGAACAAAAGATCTCGGCTTTGGTGTGTCGACGTAGCATGGTT 906			
Qy 384 HisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSer 403			
Db 907 CATCCTCGCAACAGGCTATTTCAGTTTGAGATCTTTGTCTGAAGCTTCCAAAATACGCATCA 966			
Qy 404 ValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSer 423			
Db 967 GTCATCGCTAATATATAAACATGAGACCACTACTTCTCTCACG-----AGACACATC 1020			
Qy 424 AlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGln 443			
Db 1021 AACACCAATATTTCAAGACAAGCTTGGGATATCTTTATGCTGCCACCAAGGAAGAAACGACAG 1080			

QY 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
 DB 67 AGTGTGATGATCTCGAATTCGCTGTTGTTGGCTCTGCTGTTAGCTGGCG 126
 QY 124 AlaGluAlaLysLeuGlyLeuValGlyLeuIleGlyProAspLeuProPheThr 143
 DB 127 GCTGAATCAGCTAAGTTAGGACTTAAGTTGGACTGATGGCTGACCTCTCTTCACT 186
 QY 144 AsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
 DB 187 AACACTACGGTGTGGGAAGATGAGTCAACGATCTGGCTTGCAAAATGATTCAG 246
 QY 164 HisAlaTrpLysAspThrIleValTrpLeuAspAsnAlaProValLeuIleGlyArg 183
 DB 247 CATGTTTGGAGAGATACCTTGTGTATCTGGACGATGACATCTTATACCATTTGGTGGT 306
 QY 184 AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu 203
 DB 307 GCTTATGGAAGATGATGAGTTCGAGCTTTACTTCACGAGGATCTTCAGGAGGTGTGGAG 366
 QY 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
 DB 367 TCAGGTGCTCTATCTTACCTCCAAAGTTGAGAGCATAACAGAGCTCTCTGATGGCCT 426
 QY 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
 DB 427 AGGCTGTGTTCTCTGTGAACAAACACCTTGTTCGTCGAGGCTTGCACCTGTGTCTCT 486
 QY 244 GlyAlaAlaSerGlyLysLeuGluGluTyrGluValGlyGlyProArgValCysValGln 263
 DB 487 GGACGAGCTTCGGAGAGCTCTTCAATACGAAGTTGGAGGCTAGAGTCTGTGCCAA 546
 QY 264 ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283
 DB 547 ACTGCTTACGGCTTGGAGTTGAGTGGTGAAGAGTCCATATGATCCAGAGAGTGGTG 606
 QY 284 PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluTyrPro 303
 DB 607 TTCATGATACAGAGATTATACAAACGAGAGAAATCCGGAGCTTAGAAGCTGAATATCCA 666
 QY 304 ThrPheLeuTyrValMetProMetSerProThrArgLeuPhePheGluGluThrCysLeu 323
 DB 667 ACCTTCTCTACGCCATGCTATGACAAAGACCAGAGCTCTTTGAGGAGACATGCTT 726
 QY 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
 DB 727 GCTTCAAGAGATGTCATGCCCTTGTATTTGCTTAAAGAGAGCTCTTGTGAGATTAGAG 786
 QY 344 ThrLeuGlyIleGlnValThrLysValTyrGluGluTyrSerTyrIleProValGly 363
 DB 787 ACACCTCGGAATCCGAATATAAAGACTTACGAGAGGAGTGTCTTATATCCAGTAGGT 846
 QY 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetVal 383
 DB 847 GGTCTCTGCCAAACACGGAACAAAGAAATCTCGCTTGTGTGCTGACGATGATGGT 906
 QY 384 HisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSer 403
 DB 907 CATCTCGCAACAGGCTATCTAGTTGTGAGATCCCTGTGTAAGCTCCAAATAGCATCA 966
 QY 404 ValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSer 423
 DB 967 GTCATCGCTAATATACTAAACATGAGCACCTACTTCTCTCACC-----AGACACATC 1020
 QY 424 AlaValAsnIleSerMetGlnAlaTyrSerLeuTyrProLysGluArgLysArgGln 443
 DB 1021 AACACCAATATTCAGACAGCTTGGGATACCTTATGGCCACAGAAAGAAACACAG 1080
 QY 444 ArgAlaPhePheLeu 448
 DB 1081 AGAGCATCTTCTTA 1095
 RESULT 12

AAV19123
 ID AAV19123 standard; cDNA; 960 BP.
 XX
 AC AAV19123;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Nucleotide sequence of epsilon cyclase cDNA clone 9-4.
 XX
 KW Epsilon cyclase; SSU/crtB fusion protein; carotenoid; lycopene; lutein;
 KW zeaxanthin; canthaxanthin; vitamin; colourant; ss.
 XX
 OS Brassica napus.
 XX
 PN WO9806862-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 08-AUG-1997; 97WO-US14035.
 XX
 PR 09-AUG-1996; 96US-0024145.
 XX
 PA (CALJ) CALGENE INC.
 XX
 PI Shewmaker CK;
 XX
 DR WPI; 1998-159551/14.
 XX
 PT Use of constructs comprising a carotenoid biosynthesis gene - for
 PT producing plants and seeds having altered carotenoid levels,
 PT modified fatty acid compositions or altered tocopherol levels.
 XX
 PS Example 1B; Fig 9; 70pp; English.
 XX
 CC This is the nucleotide sequence of the Brassica napus epsilon cyclase
 CC cDNA clone 9-4. It was used in the method of the invention with
 CC the SSU/crtB fusion protein (AAV19123), to create plants and seeds with
 CC altered carotenoid levels. They can be used for the production of
 CC particular carotenoids, e.g. alpha-carotene, beta-carotene, lycopene,
 CC lutein, zeaxanthin, canthaxanthin, which can be used as supplements,
 CC particularly vitamin supplements, as vegetable oil based food products
 CC and food ingredients, as feed additives in animal feeds, as colourants
 CC or therapeutic agents.
 XX

SQ Sequence 960 BP; 246 A; 203 C; 247 G; 264 T; 0 other;
 Alignment Scores:

Pred. No.:	4,03e-118	Length:	960
Score:	1202.00	Matches:	224
Percent Similarity:	88.38%	Conservative:	27
Best Local Similarity:	78.87%	Mismatches:	31
Query Match:	43.82%	Indels:	2
DB:	19	Gaps:	1

US-09-701-395A-23 (1-529) x AAV19123 (1-960)

QY 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
 DB 66 AGTGTGATGATATCGCAATTCGGCTGTTGTTGGTCTGCTGGTGGTTAGCTTGGCT 125
 QY 124 AlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr 143
 DB 126 GCAGATCAGCAGGTTAGGCTCAAGTTCGACTTCATTTGCTGATCTCTTCACT 185
 QY 144 AsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
 DB 186 AACAACTACGGTGTGGGAAGATGAGTTCACGATCTTGGCTGCAAAATGATTCAG 245
 QY 164 HisAlaTrpLysAspThrIleValTrpLeuAspAsnAlaProValLeuIleGlyArg 183
 DB 246 CATGTTTGGAGATACCTTGTGTATCTGACGATGACATCTTATACCATTTGGTCTG 305
 QY 184 AlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLysArgCysValGlu 203

```

Db 306 GCTATTGGAGAGTTAGTCACGCTTTACTTCACGAGAGCTCTGAGAGGTGTGTGGAG 365
Qy 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
Db 366 TCAGGTGTCTCGTATCTTAGCTCCAAAGTTGAGAGCATACAGAAGCTCTGATGGCTT 425
Qy 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
Db 426 AGGCTTGTTCCTGTCGAACAAACACCCCTTGTTCCTGTCAGAGCTTGCACCTGTGTCT 485
Qy 244 GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGln 263
Db 486 GGAGCAGCTTCTGGGAAGCTCTTGCATACGAAGTTGGAGGGCTAGAGTCTGTGTCCAA 545
Qy 264 ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283
Db 546 ACTGCTTAGCTGTGGAGTTGAGGTGGAAAGAGTCCATATGATCCAGAGCAGATGGTG 605
Qy 284 PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrPro 303
Db 606 TTCATGGATTACAGAGATTATACAAACGAGAAATCCGGAGCTTAGAAGCTGAATATCCA 665
Qy 304 ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu 323
Db 666 ACCTTCTCTACCCATGCTATGACAAAGACAGAGCTCTTCTTGGAGGAGATGTCTT 725
Qy 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
Db 726 GCTTCAAAAGATCTATGCTTGTGTTGCTTAAAGAGCTCTTGTGAGATTAGAG 785
Qy 344 ThrLeuGlyIleGlnValThrLysValTyrGluGluGluTyrPheSerTyrIleProValGly 363
Db 786 ACACCTCGGAATCGAATACCTAAAGACTTACGAAGAGGAATGGTCTTATATATCCAGTAGT 845
Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetVal 383
Db 846 GGTTCCTTCCCAACACGGAACAAAGAAATCTCGCCTTTTGGCGCTGCAGCTAGCATGGTA 905
Qy 384 HisProAlaThr 387
Db 906 CATCCGCAACA 917

RESULT 13
ID AAX25064
AC AAX25064 standard; DNA; 962 BP.
XX
XX AAX25064;
XX
XX 05-JUL-1999 (first entry)
XX
XX Brassica napus lycopene epsilon cyclase cDNA clone 9-4.
XX
XX Lycopene epsilon cyclase; phytoene synthase; crtB gene; carotenoid;
XX xanthophyll; pigment; beta-carotene; alpha-carotene; phytoene;
XX transgenic plant; oilseed plant; seed oil; antisense; ss.
XX
XX Brassica napus.
XX
XX WO9907867-A1.
XX
XX 18-FEB-1999.
XX
XX 06-AUG-1998; 98WO-US16466.
XX
XX 08-AUG-1997; 97US-0908758.
XX
XX (CALJ ) CALGENE LLC.
XX
XX Shewmaker CK;
XX
XX WPI; 1999-180495/15.
XX

```

```

PT Altering xanthophyll content of seeds by transformation - used to
PT produce seed oils of increased carotenoid content, e.g. Brassica and
PT cotton
XX
XX Example 1; Fig 9; 92pp; English.
XX
CC This is the nucleotide sequence of Brassica napus lycopene
CC epsilon cyclase cDNA clone 9-4, which was obtained by PCR using
CC primers designed from an Arabidopsis epsilon cyclase gene. The
CC xanthophyll content of seeds is altered by transforming cells of a
CC host plant with at least one construct containing: (i) a transcription
CC initiation region of a gene preferentially expressed in seeds; (ii)
CC the sequence for a plastid transit peptide; (iii) DNA from a
CC carotenoid synthesis gene coding region, and (iv) a transcription
CC terminator. The transformed cells are regenerated to plants and
CC these, or their progeny, grown to produce seeds. The method is
CC especially used to increase the carotenoid content in oilseed
CC plants. Expression of phytoene synthase, encoded by the crtB gene
CC (see AAX25063), together with an antisense epsilon cyclase gene
CC results in large increases in levels of alpha-carotene,
CC beta-carotene and phytoene, and an alteration of alpha- to
CC beta-carotene ratio, in seeds of transformed plants.
XX
SQ Sequence 962 BP; 247 A; 203 C; 248 G; 264 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 4.04e-118 Length: 962
Score: 1202.00 Matches: 224
Percent Similarity: 88.38% Conservativeness: 27
Best Local Similarity: 78.87% Mismatches: 31
Query Match: 43.82% Indels: 2
DB: 20 Gaps: 1

```

US-09-701-395a-23 (1-529) x AAX25064 (1-962)

```

Qy 106 SerValMetAspLeu-----valValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
Db 66 AGTGTGATGGATATCTGCAGAAATTCGGCTTGTGTGTCTCTGCTGTGGTGTAGCTTGGCT 125
Qy 124 AlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr 143
Db 126 GCAGAAATCAGCAAGGTAGGTCTCAAGTTGAGTCACTATGTGCTCTGATCTCTCTTCACT 185
Qy 144 AsnAsnTyrGlyValTyrGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
Db 186 AACAACTACGGTGTGGGAAGATGAGTTCAACGATCTTGGCTTGCAGAAATGATTAGAG 245
Qy 164 HisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArg 183
Db 246 CATGTTTGGAGATACCTTGTGTATCTGTGACGATGACAAATCTATTACCATTTGGTCTGT 305
Qy 184 AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu 203
Db 306 GCTTATGGAGAGTTAGTCGAGCTTACTTCCAGGAGAGCTCTTGGAGAGGTGTGTGGAG 365
Qy 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
Db 366 TCAGGTGTCTCGTATCTTAGCTCCAAAGTTGAGAGCATAACAGAAGCTCTGTATGGCTT 425
Qy 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
Db 426 AGGCTTGTTCCTGTCGAACAAACACCCCTTGTTCCTGTCAGAGCTTGCACCTGTCTCT 485
Qy 244 GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGln 263
Db 486 GGAGCAGCTTCTGGGAAGCTCTTGCATACGAAGTTGGAGGGCTAGAGTCTGTGTCCAA 545
Qy 264 ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283
Db 546 ACTGCTTAGCTGTGGAGTTGAGGTGGAAAGAGTCCATATGATCCAGAGCAGATGGTG 605
Qy 284 PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrPro 303

```

Db 606 TTCATGGATTACAGAGATTATACAAACGAGAAATCCGGAGCTTAGAAGCTGAATATCCA 665
 Qy 304 ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluThrCysLeu 323
 Db 666 AGCTTCTCTAGCCATGCTATGACAAAGACAGAGCTTCTTTGAGGAGACATGCTT 725
 Qy 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
 Db 726 GCTTCAAAAGATGTCATGCCCTTTGATTGCTTAAAGAGAGCTCTTGTGAGATTAG 785
 Qy 344 ThrLeuGlyIleGlnValThrLysValTyrGluGluTyrSerTyrIleProValGly 363
 Db 786 ACACCTCGAATCCGAACTAAAGACTTACGAAGAGGATGCTATATCCAGTAGGT 845
 Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal 383
 Db 846 GGTTCCTTCCCAACACGGAACAAAGAAATCTCGCTTTGGCGTGCAGCTAGCATGTA 905
 Qy 384 HisProAlaThr 387
 Db 906 CATCCCGCAACA 917

RESULT 14

ABA97355

ID ABA97355 standard; cDNA; 962 BP.

AC ABA97355;

DT 12-APR-2002 (first entry)

DE Brassica napus epsilon cyclase cDNA clone 9-4.

KW Carotenoid compound; oilseed; plant; vitamin A; ss; gene.

OS Brassica napus.

PN WO200188169-A2.

PD 22-NOV-2001.

PF 11-MAY-2001; 2001WO-US15264.

PR 12-MAY-2000; 2000US-0570140.

PX (MONS) MONSANTO TECHNOLOGY LLC.

PY Shewmaker CK;

PZ WPI; 2002-075318/10.

PT Altering carotenoid level/content in oilseed plants, by transforming
 PT plants with construct having DNA sequence encoding a carotenoid
 PT biosynthesis gene, plastid transit peptide, transcription initiator and
 PT terminator

PS Disclosure; Page 71; 114pp; English.

CC This invention relates to altering the carotenoid content or
 CC composition in seed from a host plant. This is achieved by transforming
 CC cells of host plant with a construct of operably linked components, a
 CC transcriptional initiation region from a gene expressed in a plant seed,
 CC a plastid transit peptide, a nucleic acid sequence encoding a carotenoid
 CC biosynthesis gene from eukaryotic source, and a transcriptional
 CC termination region. The method is useful for increasing carotenoid
 CC content in seed and endosperm of a host plant and altering
 CC carotenoid composition in a plant seed. The method is also useful for
 CC screening transformed corn seeds or transformed endosperms, where the
 CC transformed seed and endosperms are visually determined and selected
 CC based on yellow, orange or red colour as result of the increased
 CC carotenoid content. The transformed seeds provide a source of
 CC modified oils and the oil extracted from the seeds is useful as a food
 CC colourant, or as a food oil with high alpha and beta-carotene levels for
 CC prevention of vitamin A deficiency which can result in night blindness.

CC The level of lutein is also increased in seeds of transformed plants.
 CC The modified oil obtained from the transformed seeds are more oxidatively
 CC stable than naturally occurring oils and are low saturate, high oleic and
 CC low linolenic. This sequence represents the nucleotide sequence of
 CC the Brassica napus epsilon cyclase cDNA clone 9-4.

XX SQ Sequence 962 BP; 247 A; 203 C; 248 G; 264 T; 0 other;

Alignment Scores:

Pred. No.: 4,04e-118 Length: 962
 Score: 1202.00 Matches: 224
 Percent Similarity: 88.38% Conservatives: 27
 Best Local Similarity: 78.87% Mismatches: 31
 Query Match: 43.82% Indels: 2
 DB: 24 Gaps: 1

US-09-701-395A-23 (1-529) x ABA97355 (1-962)

Qy 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
 Db 66 AGTGTGATGGATATCTGCAGAAATTCGGCTTGTGCTGCTGCTGCTTAGCTTGGCT 125
 Qy 124 AlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr 143
 Db 126 GCAGAAATCAGCAAGGTTAGGTCTCAAGTTGGACTTCATTGGTCTGCTTCTTCTACT 185
 Qy 144 AsnAsnTyrGlyValTyrPheLysAspLeuGlyLeuGluArgCysIleGlu 163
 Db 186 AACAACTACGGTGTGGGAAGATGAGTTCAACGATCTTGGCTTGCAGAAATGATTGAG 245
 Qy 164 HisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArg 183
 Db 246 CATGTTTGGAGAGATACCTTGTGTATCTGGACGATGACAATCCTATTACCATTTGGCT 305
 Qy 184 AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu 203
 Db 306 GCTTATGGAGAGTAGTAGTACGAGCTTACTTCCAGGAGAGCTTCTTGAGAGGTGTGTGAG 365
 Qy 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
 Db 366 TCAGTGTCTCGTATCTTAGCTCCAAAGTTGAGAGCATACAGAGCTTCTTGATGGCCTT 425
 Qy 224 SerLeuValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
 Db 426 AGGCTTGTCTCTGTGAACAAACACCCCTTGTCTGTCAGGCTTGCCACATGCTTCT 485
 Qy 244 GlyAlaAlaSerGlyLysLeuGluTyrGlyValGlyGlyProArgValCysValGln 263
 Db 486 GGACGAGCTTCTGGGAAGCTCTTGCAATACGAAGTTGGAGGCTTAGAGTCTGTGTCAC 545
 Qy 264 ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283
 Db 546 ACTGCTTACGGCTTGGAGTTGAGTGGAAAGAGCTCCATATGATCCAGAGCAGATGGTG 605
 Qy 284 PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrPro 303
 Db 606 TTCATGGATTACAGAGATTATACAAACGAGAAATCCGGAGCTTAGAAGCTGAATATCCA 665
 Qy 304 ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu 323
 Db 666 ACGTTTCTTACGCCATGCCCTATGACAAAGACACAGAGTCTTCTTTGAGGAGACATGCTT 725
 Qy 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
 Db 726 GCTTCAAAAGATGTCATGCCCTTTGATTGCTTAAAGAGAGCTTCTTGTGAGATTAG 785
 Qy 344 ThrLeuGlyIleGlnValThrLysValTyrGluGluTyrSerTyrIleProValGly 363
 Db 786 ACACCTCGAATCCGAACTAAAGACTTACGAAGAGGATGCTTATATCCAGTAGGT 845
 Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal 383
 Db 846 GGTTCCTTCCCAACACGGAACAAAGAAATCTCGCTTTGGCGTGCAGCTAGCATGTA 905

QY 384 HisProAlaThr 387
|||||
Db 906 CATCCGCAACA 917
RESULT 15
AAC44887
ID AAC44887 standard; DNA; 1039 BP.
XX AC AAC44887;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 44503.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0121180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

```

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157885.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment Scores:
Pred. No.: 1.8e-116
Score: 1187.00
Percent Similarity: 79.01%
Best Local Similarity: 79.01%
Query Match: 68.51%
DB: 43.27%
Indels: 17
Gaps: 5

```

US-09-701-395A-23 (1-529) x AAC44887 (1-1039)

QY 1 MetGluLeuGlyValArgAsnLeu-----IleSerSerCysProValTrp 16

```

Db 26 ATGGAGTGTGTGGGCTAGGAATTCGACGAATGGCGGTTCACATTCCTCATGG 85
QY 17 ThrPheGlyThrArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGly 36
Db 86 AGTTCTCGAAGGAATTTCCAGTGGTTAAGAGATACAGCTATAGGAATATTCGTTCCGT 145
QY 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerArg 54
Db 146 ---TTGTGT-----AGTGTACAGCTAGCGCGCGGGAAGTTCGCTAGTAGAG 190
QY 55 SerSerValAlaTyrLysGluGlyPheValAspGluAspPheIleLysAlaGlyGly 74
Db 191 AGTTGTGTAGCGGTGAGAGAGATTTCCGCTGACGAGAGAGATTTGTGAAAGCTGGTGT 250
QY 75 SerGluLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
Db 251 TCTGAGATTCTATTGTTCAAATGCAGCAGACAAAGATATGGATGAACAGCTTAGCTT 310
QY 95 AlaAspLysLeuProTrieProPheGlyGluSerValMetAspLeuValIleGly 114
Db 311 GTTGATAAGTTGCCCTATATCAATTTGGTGTGCTTGGANCTAGCTAGGTATTGT 370
QY 115 CysGlyProAlaGlyLeuSerLeuAlaLagLuuAlaLysLeuGlyLeuLysValGly 134
Db 371 TGTGTCTCTGCTGTTAGCTTGGCTGCAGANTCAGCTAAGCTTGGATTGAAGTTGGA 430
QY 135 LeuIleGlyProAspLeuProPheThrAsnTyrGlyValTrpGluAspGluPheLys 154
Db 431 CTCATTGGTCCAGATCTCTCTTTACTAACAAATACGGTGTTCGGGAAGATGAATCAAT 490
QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAsp 174
Db 491 GATCTTGGGCTGCAAAATGATTGAGCATGTTGGAGAGACATATGTGTATCTGGAT 550
QY 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHis 194
Db 551 GATGACAAGCCTATTACCATGGCGGTCTATGGAAGAGTTAGTCACGCTTTGCTCCAT 610
QY 195 GluGluLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGlu 214
Db 611 GAGGAGCTTTTGAGGAGGTGTGCGAGTCAAGTGTCTGCTACCTTAGCTCGAAAGTTGAC 670
QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle 234
Db 671 AGCATACAGAACTTCTGATGCCTTAGCTTGTGTGTGACGACAAATAACGTCATT 730
QY 235 ProCysArgLeuAlaThrValAlaSerGlyAlaLaserGlyLysLeuGluTrpGlu 254
Db 731 CCTGCGAGCTTGCCACTTGTCTTCTGAGCAGCTTCGGGAAAGCTTTCGAATACGAA 790
QY 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsn 274
Db 791 GTTGTGGACCTAGAGTCTGTGTCGCAACTGCATACGCGGTGGAGGTGAGTGGAAAT 850
QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlyLys 294
Db 851 AGTCCATATGATCCAGATCAAAATGTTTTCATGATTACAGAGATTATATACGAGAAA 910
QY 295 LeuGlnCysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThr 314
Db 911 GTTCGGAGCTTAGAAGCTGAGTATCCAAAGCTTCTGTGATGCCCATGCTATGACAAAGTCA 970
QY 315 ArgLeuPhePheGlu-----GluThrCysLeuAlaSerLysAspAlaMe 329
Db 971 AGACTCTTCTGAGCTGGAAATCCTTTTTCAGGAGACATGTTTGGCCTCAAAAGATGTCAT 1030
QY 329 tProPhe 331
Db 1031 GCCCTTT 1037

```

Search completed: May 21, 2003, 22:28:53

Thu May '22 09:39:16 2003

us-09-701-395a-23.rng

Page 19

Job time : 396 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:12:43 ; Search time 4278 Seconds
(without alignments)
3598.734 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRNLISCPVWTFGT.....VRHLLSDPSGAVNVRAYLER 529

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cn2.1/USPTO.spool/US09701395/runat_15052003.165750.3330/app_query.fasta_1.711
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09701395.ecgn.1.1.2586@runat_15052003.165750.3330 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.mu.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2743	100.0	1898	8	AF321536 Adonis pa
2	2728	99.5	1848	8	AF321535 Adonis pa
3	1966.5	71.7	1606	8	AY079371 Arabidops
4	1966.5	71.7	1836	8	AY040024 Arabidops
5	1959.5	71.4	1860	8	AR003634 Sequence
6	1959.5	71.4	1860	8	ATU50738 Arabodopsi
7	1934	70.5	1780	8	AF321538 Lactuca s
8	1934	70.5	1830	8	AF251016 Tagetes e
9	1934	70.5	1887	6	AR152414 Sequence
10	1913.5	69.8	1916	8	AY099485 Tagetes e
11	1909.5	69.6	1697	8	AY099485 Tagetes e
12	1892	69.0	2357	8	AF463497 Spinacia
13	1786	55.1	1550	8	AF486650 Citrus x
14	1633.5	59.6	3634	8	AF117257 Arabidops
15	1633.5	59.6	78423	8	AB024035 Arabidops
16	1518.5	55.4	1378	8	AF321537 Solanum t
17	1394	50.8	1272	6	AX350852 Sequence
18	1349	49.2	144879	8	AP003332 Oryza sat
19	1202	43.8	962	6	AX350851 Sequence
20	1113	40.6	787	8	AF450280 Citrus si
21	874	31.9	605	8	AF229684 Daucus ca
22	783.5	28.6	2027	8	AF240787 Citrus si
23	778.5	28.4	1830	8	AF152246 Citrus x
24	778.5	28.4	3313	8	AY094582 Citrus si
25	775.5	28.3	2026	8	AF321534 Adonis pa
26	761	27.7	1785	8	X86221 C.annuum mR
27	761	27.7	1942	6	A58419 Sequence 1
28	755	27.5	1906	8	AF251017 Tagetes e
29	750.5	27.4	1650	6	AR022630 Sequence
30	750.5	27.3	1959	6	AR152412 Sequence
31	749.5	27.3	1650	8	LECRTL1
32	748	27.3	1537	8	AY091396 Arabidops
33	748	27.3	1946	8	ATU50739 Arabidops
34	748	27.3	1975	8	AY059749 Arabidops
35	748	27.3	2814	8	AF117256 Arabidops
36	748	27.3	87937	8	ATAC009400 Arabidops
37	747	27.2	1471	8	AY099484 Tagetes e
38	747	27.2	1661	8	ATHLYC
39	739	26.9	1590	6	AX350853 Sequence
40	737	26.9	1614	6	AR022629 Sequence
41	737	26.9	1614	8	NTCRTL1
42	736	26.8	497	8	AF212130 Daucus ca
43	729	26.6	1696	8	AF489520 Sanderson
44	719.5	26.2	1666	8	AF254793 Lycopersi
45	714.5	26.0	1498	8	STU272136 Solanum t

ALIGNMENTS

RESULT 1

AF321536 1898 bp mRNA linear PLN 28-FEB-2001
 LOCUS Adonis palaestina lycopene epsilon-cyclase mRNA, complete cds.
 DEFINITION AF321536
 ACCESSION AF321536
 VERSION AF321536.1 GI:12746308
 KEYWORDS
 SOURCE Adonis palaestina.
 ORGANISM Adonis palaestina.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales; Ranunculaceae; Adonis.
 REFERENCE 1 (bases 1 to 1898).
 AUTHORS Cunningham, F. X. Jr. and Gantt, E.
 TITLE One ring or two? Determination of ring number in carotenoids by lycopene varepsilon-cyclases
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (5), 2905-2910 (2001)
 MEDLINE 21126984
 PUBMED 11226339
 REFERENCE 2 (bases 1 to 1898)
 AUTHORS Cunningham, F. X. Jr.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2000) Cell Biology and Molecular Genetics, University of Maryland, Microbiology Building, Campus Drive, College Park, MD 20742, USA
 FEATURES
 source Location/Qualifiers
 1. .1898
 /organism="Adonis palaestina"
 /db_xref="taxon:151078"
 /tissue_type="Immature and developing flower buds"
 113. .1702
 /function="converts lycopene to delta-carotene and epsilon-carotene"
 /note="evidence by complementation in lycopene-accumulating Escherichia coli"
 /codon_start=1
 /evidence=experimental
 /product="lycopene epsilon-cyclase"
 /protein_id="AAK07432.1"
 /db_xref="GI:12746309"
 /translation="MELGVRLNLSVCPVWTFGRNLSKSLVNIHRYGSSCRVDFQVRADGSGSRVAVYKEGFVDEDFIKAGSELFLVQMQTKSMERKAKLADLPPIIPFGESVMDLAVIIGCGPAGLSLAEEAKLGLKVLGPDLPFTNNYGVWEDFKDLGLERCIHAWKDTVYLDNDAPVLIGRAIVRSVRLHEELKRCVSGSYLDSKVERITEAGDSHVTVCNEFIIPCLRLATVAGSAAKLLLEYEYGVGPRVCVQYAYGVEVENNPYDNLNMFMDRTFMQQLQCESEETFLYVMPMSPTLFFETCTCLASDKAMPFDLLKRLMSRLKLTGIQVTKYEEWSEYIPVGSGLPNTKONLAFGAASAMVHPATGYSVVRSLSAPKYSVIAKILKQNSAVYVSGSSAVNISMQWSSLPKPKRQKRAFFLVGLLEIVQIDIEATRTFFTRFLPTMMWNGFLGSLSLSDFLVLFMTMFVLPNSMRMSLVRLHLLSDPSGAVMVRAYLER"
 BASE COUNT 523 a 342 c 437 g 596 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.25e-226 Length: 1898
 Score: 2743.00 Matches: 529
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-09-701-395A-23 (1-529) x AF321536 (1-1898)
 QY 1 MetGluLeuLeuGlyValArgAsnLeuIleSerSerCysProValTrpThrPheGlyThr 20
 DB 113 ATGGAACACTTGGTGTTCACACCTCATCTCTCTGCGCCCTGTGGACTTTTGAACA 172
 QY 21 ArgAsnLeuSerSerSertysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArg 40
 DB 173 AGAAACCTTAGTAGTTCAAACCTAGCTTATACATACATACATGATGCTTCTTCTGAGA 232
 QY 41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerSerValAlaTyrLys 60
 DB 233 GTAGATTTTCACTGAGAGCTGATGTGGAAGCGGAGTAGAAGTCTCTGCTTATAA 292

QY 61 GluGlyPheValAspGluGluAspPheIleLysAlaGlySerGluLeuLeuPheVal 80
 DB 293 GAGGTTTGTGGATGAAGAGGATTTATCAAGCTGGTGGTCTGAGCTTTGTGTC 352
 QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
 DB 353 CAAATGACGACAAACAAGTCTATGAGAACAGCCAGCTGCCGATAGTTGCCACA 412
 QY 101 IleProPheGlyCysSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
 DB 413 ATACCTTTGGAGAAATCCGTGATGGACTTGGTGTATAGTGTGTGACCTGCTGCTT 472
 QY 121 SerLeuAlaAlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
 DB 473 TCACTGGCTGCAGAGCTGCTAAGCTAGGTTGAAAGTTGGCTTATGGTCTGATCTT 532
 QY 141 ProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGluArg 160
 DB 533 CTTTTACAAATATATATGGTGTGTGGAGACGAGTTCAAAGATCTTGGACTTGAACGT 592
 QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180
 DB 593 TGTATCGAGCATGCTTGGAGGACACCATCTATATCTTGATATGATGCTCTGCTGCTT 652
 QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArg 200
 DB 653 ATGTGCTGTGATATGGACGAGTTAGTCGACATTTGCTACATGAGGAGTCTCTGAAAAG 712
 QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
 DB 713 TGTGTGGAGTCAGGTGATCATCTCTGGATTCTAAGTGGAAAGATCATCTGAAGCTGGT 772
 QY 221 AspGlyHisSerLeuValCysGluAsnIlePheIleProCysArgLeuAlaThr 240
 DB 773 GATGCCATAGCTTGTAGTTGTGAAATGAGATCTTTATCCCTTCGAGCTTGCTACT 832
 QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuIleThrGluValGlyProArgVal 260
 DB 833 GTTCATCTCGAGCAGCTTCAGGGAACCTTTGGAGTATGAAGTAGTGCCCTCGTGT 892
 QY 261 CysValGlnThrAlaTyrGlyValGluValGluAsnAsnProTyrAspProAsn 280
 DB 893 TGTGTCAAAACCCCTTATGGGGTGGAGTTGAGTGGAGAACATCATACGATCCCAAC 952
 QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnLysLeuGlnCysSerGluGlu 300
 DB 953 TTAATGTTATTCATGACTACAGAGACTATATCAACAGAAATACAGTCTCGGAAGAA 1012
 QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu 320
 DB 1013 GAATATCCAACATTTCTCTATGTCATGCCCATGTGCCCAACAGACTTTTTTTTGAGAA 1072
 QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLysArgLysLeuMetSer 340
 DB 1073 ACCTGTTTGGCTCAAAAGATGCCATGCCATTCGATCTACTCAAGAGAAACTGATGCA 1132
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTrpSerTyrIle 360
 DB 1133 CGATTGAGACTCTGGTATCCAAAGTTTATGAAGAGGATGGTCATATATT 1192
 QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAla 380
 DB 1193 CTTGTTGGTGTCTTTTACCAACACAGAGCAAAAGAACCTAGCATTTGGTGTGACGA 1252
 QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
 DB 1253 AGCATGGTGCATCCAGCAACAGGCTATTCGGTTGTACGGTCCTCCTCAGAAAGCTCCA 1312
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
 DB 1313 TATGCTTCTGTAATGCAAGATTTTGAAGCAAGATAACTCTCGGTATGTGGTTCTTGGGA 1372

```

QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
Db 1373 CAAAGTAGTCAGTAACATTTCAATGCAAGCAGTGGAGCAGTCTTTGGCCAAAGGACGGA 1432
QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
Db 1433 AAACGTCAAGAGCATCTCTCTTTGGATTAGAGCTTATTGTGCAGCTAGATATTGAA 1492
QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480
Db 1493 GCAACAGAACATCTTTAGAACCTTCTCCGCTTCCCACTTGGATGGTGGGGTTTC 1552
QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
Db 1553 CTGGGTCTTCACTATCATCTTTTCGATCTCGTCTGTTTTCATGATGTATGTTTGTG 1612
QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
Db 1613 GCGCCAAACAGCATGAGGATGTCACCTGTGAGACATTTGCTTTCAGATCCTTCTGTGCA 1672
QY 521 ValMetValArgAlaTrpLeuGluArg 529
Db 1673 GTTATGGTAAGACTTACCTCGAAAGG 1699

RESULT 2
AF321535 1848 bp mRNA linear PLN 28-FEB-2001
LOCUS AF321535
DEFINITION Adonis palaestina lycopene epsilon-cyclase mRNA, complete cds.
ACCESSION AF321535
VERSION AF321535.1 GI:12746306
KEYWORDS
SOURCE Adonis palaestina.
ORGANISM Adonis palaestina.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Adonis.
REFERENCE 1 (bases 1 to 1848)
AUTHORS Cunningham, F.X. Jr. and Gantt, E.
TITLE One ring or two? Determination of ring number in carotenoids by
lycopene varepsilon-cyclases
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (5), 2905-2910 (2001)
MEDLINE 21126984
PUBMED 11226339
REFERENCE 2 (bases 1 to 1848)
AUTHORS Cunningham, F.X. Jr.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2000) Cell Biology and Molecular Genetics,
University of Maryland, Microbiology Building, Campus Drive,
College Park, MD 20742, USA
FEATURES
Location/Qualifiers
1..1848
/organism="Adonis palaestina"
/db_xref="taxon:151078"
/tissue_type="Immature and developing flower buds"
116..1705
/function="converts lycopene to delta-carotene and
epsilon-carotene"
/notes="evidence by complementation in
lycopene-accumulating Escherichia coli"
/codon_start=1
/evidence-experimental
/product="lycopene epsilon-cyclase"
/protein_id="AAK07431.1"
/db_xref="GI:12746307"
/translation="MELIGVRNLISSCPWTGTRNLSSKLAYNHRVSGSRVDFQ
VRADGSGRTSVAKGEFVDEDFIKAGSSELFVQMOTKSMERQAKLADKLPIF
FGESYMDLVITGCGPAGLSLAAEAAKLGKLVGLIGDLPFTNNYGVWDEFKDLGLER
CTEAWKDTIVYLDNDAPVLIIGRAYGRVSRHLHLEELKRCVSGVSYLNSKVERITE
ADGHSVLVCENDIEIPCRLATVASGAASGKLLEYVGGPRVCVOTAYGVEVEVNNP
YDPLMFMVDYRDMQKLOCSEEEYPTFLYVMPSPFLFEETCLASKDAMPDLL
KRKLMSRLKTLGIQVTKYEEEWSPYVGGSLPNTQKNLAFGAASVMPHATGYSVV
RSLSEAPYASVIAKILKODNSAYVVGSSAVNTSMQASWSLWPKRKRQRAFLFG
LELIVOLDIEATRTFTFRFLPTWWMGFLGSSLSFDLVLFSMYFVLAPNSMRMS

```

```

BASE COUNT 522 a 334 c 420 g 572 t
ORIGIN
Alignment Scores: 1.54e-224 Length: 1848
Pred. No.: 2728.00 Matches: 524
Score: 100.00% Conservative: 5
Percent Similarity: 99.05% Mismatches: 0
Best Local Similarity: 99.45% Indels: 0
Query Match: 8 Gaps: 0
DB:
US-09-701-395a-23 (1-529) x AF321535 (1-1848)
QY 1 MetGluLeuLeuGlyValAlaArgAsnLeuIleSerSerCysProValTrpThrPheGlyThr 20
Db 116 ATGGAACCTACTTGTGTGTCGCAACCTCATCTCTTCTGCCCTGCTGGACCTTTTGGAA 175
QY 21 ArgAsnLeuSerSerSerLysLeuAlaTrpAsnIleHisArgTyrGlySerSerCysArg 40
Db 176 AGAACCTTAGTACTTCAAAACTAGCTTATAACATACATCATGATGTTCTTCTTGTA 235
QY 41 ValAspPheGlnValArgAlaAspGlySerGlySerArgSerSerValAlaTrpLys 60
Db 236 GTAGATTTTCAAGTGAAGGCTGTGTTGGAAGCGGAGTAGAACCTCTCTTCTTATAAA 295
QY 61 GluGlyPheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheVal 80
Db 296 GAGGGTTTGTGGACGAGGAGGATTTATCAAGCTGTGTGTTCTTGAGCTTTTGTGTC 355
QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
Db 356 CAATGCAGCAAAACAAAGTCTATGGAGAAACAGCCAGCTCGCGATAAGTTGCCACCA 415
QY 101 IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
Db 416 ATACCTTTTCGGAGAATCTGTGATGACTTGTGTTGTAATAGGTGTGGACCTGCTGCT 475
QY 121 SerLeuAlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
Db 476 TCACCTGGCTGCAGAGCTGTAACTAGGCTTGAAGCTTGAAGCTTATTTGGTCTGTGAT 535
QY 141 ProPheThrAsnAsnTrpGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArg 160
Db 536 CTTTTTACAAATAATTATGTTGTGGAGACAGAGTTCAAAGATCTTTGGACTTGAACGT 595
QY 161 CysIleGluHisAlaTrpLysAspThrIleValTrpLeuAspAsnAspAlaProValLeu 180
Db 596 TGTATCGAGCATGCTTGGAGGACACCATCATATATCTTGACAATGATGCTCTGCTCTT 655
QY 181 IleGlyArgAlaTrpGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArg 200
Db 656 ATTGGTCTGCATATGACGAGTTAGCCGCGATTTGTCGATGAGAGAGTCTGCTGAAAAG 715
QY 201 CysValGluSerGlyValSerTrpLeuAspSerLysValGluArgIleThrGluAlaGly 220
Db 716 TGTCTCAGTCAGCTGTATCATATCTGAATTTCTAAAGTGGAAAGAGTCACTGAAGCTGT 775
QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
Db 776 GATGGCCATAGTCTTGTAGTTTGTGAACACGACATCTTTATCCCTTTCAGGCTTCTCT 835
QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluValGlyGlyProArgVal 260
Db 836 GTTGCATCTGGAGCAGCTTTCAGGAAACTTTTGGAGTATGAAGTAGTGGCCCTCGTGT 895
QY 261 CysValGlnThrAlaTrpGlyValGluValGluValGluAsnAsnProTyrAspProAsn 280
Db 896 TGTGTCCAAACTGCTTATGTTGTGGAGTTGAGGTGGAGAGCAACATCCATACGATCCCA 955
QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300
Db 956 TTAATGTTATTTATGGACTACAGAGACTATATATGCAACAGAAATATACAGTCTCGGAAGA 1015

```

QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePheGlu 320
 Db 1016 GAATATCAACATCTCTATGTATGTCATGCCATGCGCCACAAAGACATTTTGTGAGGAA 1075
 QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer 340
 Db 1076 ACCGTGTGGGCTCAAAAGATGCCATGCTTTCATCTACTAGAGAGAACTAATGTCA 1135
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTyrSerTyrIle 360
 Db 1136 CGATTGAAGACTCGGTATCCCAAGTTACAAATTTATGAAGGAATGCTTATATT 1195
 QY 361 ProValGlyLysSerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAla 380
 Db 1196 CCTGTGGGGTCTTTACCAACACAGAGCAAAAGAACTAGATTTGTGTGTCAGCA 1255
 QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
 Db 1256 AGCATGTGTCATCCAGCAACAGGCTATTCGGTTGTAGCATCTATCAGAAAGTCCAAA 1315
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
 Db 1316 TATGCTCTCTAATGTCAAGATTTTGAAGCAAGATAACTCTCATATGTGTTCTGGA 1375
 QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
 Db 1376 CAAGCAGTGCAGTAACATTTCAATGCAAGCATGAGCAGTCTTTGGCCAAAGAGCGA 1435
 QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
 Db 1436 AAACGTCAAGAGCATCTTCTTTTCGGGTAGACCTTATGTGCGAGTACATATTTGAA 1495
 QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpGlyPhe 480
 Db 1496 GCACACAGACGTTCTTTAGAACCTTCTCCGCTTCCCACTGGATGGTGGGTTC 1555
 QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 Db 1556 CTGGGTCTTCACTATCATCTTTCGATCTTGTATGTTTTCATGATGATGTTTGTG 1615
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 Db 1616 GCCCGCAAGCATGAGGATGTCTTGTGAGCATTTGCTTTCAGATCTTCTGTCGA 1675
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 Db 1676 GTTATGTTAAAGCTTACTCTGAAAGG 1702
 RESULT 3
 LOCUS AY079371
 DEFINITION Arabidopsis thaliana putative lycopene epsilon cyclase (At5g57030)
 VERSION AY079371.1
 KEYWORDS FLI_CDNA.
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 1606)
 Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Eguu,P.,
 Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q., Chen,B.,
 Karlin-Neumann,G., Kim,C., Lam,B., Miranda,M., Nguyen,M.,
 Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and
 Theologis,A.
 TITLE Arabidopsis Open Reading Frame (ORF) Clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1606)
 Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
 Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,

Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
 Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
 Jones,T., Kamiya,A.C., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
 Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
 Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
 Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
 Direct Submission

TITLE
JOURNAL

Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL CDNAs (RAFL cDNA : RIKEN
 Arabidopsis Full-Length cDNA): Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and
 sequenced the PUNI (ORF) clones using the RAFL CDNAs: Yamada,K.,
 Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
 Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
 Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
 Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
 Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
 Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally
 to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
 /PGECC) contributed equally to this work as PIs.

FEATURES
source

1..1606
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="U10203"
 /note="This clone is in pUNI 51."
 ecotype: Columbia"
 1..1606
 /gene="At5g57030"
 /gene="At5g57030"
 /codon_start=1
 /evidence="experimental
 /product="putative lycopene epsilon cyclase"
 /protein_id="AA185102.1"
 /db_xref="GI:19310743"

gene

1..1606

CDS

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

1..1575

3'UTR

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

1576..1606

BASE COUNT

439 a 300 c 391 g 476 t

ORIGIN

Alignment Scores:

Score: 2.6e-159 Length: 1606
 Percent Similarity: 1966.50 Matches: 377
 Best Local Similarity: 81.84% Conservative: 60
 Query Match: 70.60% Mismatches: 80
 DB: 71.69% Indels: 17
 Gaps: 5

US-09-701-395A-23 (1-529) x AY079371 (1-1606)

QY 1 MetGluLeuGlyValArgAsnLeu-----IleSerSerCysProValTtp 16
 Db 1 ATGGAGTGTGTGGGCTAGGAATTTTCGACGAATGGCGGTTTCAACATTTCCGTCATGG 60

QY 17 ThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGly 36
 Db 61 AGTTGTCGAAGGAATTTCCAGTGGTTAAGAGATACAGTATAGGAATATTCGTTTCGGT 120
 QY 37 SerSerCysArgValaspPheGlnValArgAlaAspGlyGly-----SerGlySerArg 54
 Db 121 ---TTCTGT-----AGTGTACAGCTAGGCGCGCGGAAGTTCCGGTAGCTGAG 165
 QY 55 SerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
 Db 166 AGTTGTGTAGCGGTGAGAGAGATTTTCCTGACGAGAGAGATTTTGTGAAGCTGGTGGT 225
 QY 75 SerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
 Db 226 TCTGAGATCTATTCTTCAATTCACAGCAGCAACAAGATATGATGATGAACAGCTTAAGCTT 285
 QY 95 AlaAspLysLeuProIleProPheGlyGluSerValMetAspLeuValValIleGly 114
 Db 286 GTTGATAGTTGCTCTATATCAATTTGGTGTAGTGGTCTTTGGATCTAGTGGTTATGGT 345
 QY 115 CysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGly 134
 Db 346 TGTGGTCTCTGCTGTTAGCTTGGCTGCAGATCAGCTAAGCTTGGATTAAGATTGA 405
 QY 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLys 154
 Db 406 CTCATTGGTCCAGATCTCTCTTTACTAACAAATACGGTGTGGGAAGATGAATCAAT 465
 QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTrpLeuAsp 174
 Db 466 GATCTTGGCTCGCAAAATGTATGACATGTTGGAGAGAGACTATGTGTATCTGGAT 525
 QY 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHis 194
 Db 526 GATGACAAAGCTATTACCATTTGSCCGCTTATGGAAGACTTAGTCGACGTTTGCCTCAT 585
 QY 195 GluGluLeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGlu 214
 Db 586 GAGGAGCTTTTGGAGAGGTGTGTCGAGTCTGCTGACGTTTACCTAGCTCGAAAGTTGAC 645
 QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle 234
 Db 646 AGCATACAGAAAGCTTCTGATGCCCTTAGACTTGTGTGTGACCAACAATAACGTCATT 705
 QY 235 ProCysArgLeuAlaThrValAlaSerGlyValAlaSerGlyLysLeuLeuGluTyrGlu 254
 Db 706 CCTCGAGGCTGCCCACTGTGCTCTGGAGCAGCTTCGGGAAGCTCTTGCAATAGGAA 765
 QY 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsn 274
 Db 766 GTTGTGGACCTAGAGTCTGTGTCAAACTGTCATAGCGGTGAGGTTGAGGTGGAAAT 825
 QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLys 294
 Db 826 AGTCCATATGATCCAGATCAATGGTTTTCATGGATTACAGATATATCTAAGAGAAA 885
 QY 295 LeuGlnCysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThr 314
 Db 886 GTTCCGAGGCTTAGAAGCTGAGTATCCAACTGTTCTGTACGCCATGCCATGACAAAAGTCA 945
 QY 315 ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
 Db 946 AGACTCTTCTCGAGAGACATGTTGGCTCAAAAGATGTCATGCCCTTTGATTTGCTA 1005
 QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
 Db 1006 AAAACGAAGCTCATGTATAGATTAGATACACTCGGAATTCGAATTCCTAAGACTTAGGAA 1065
 QY 355 GluGluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeu 374
 Db 1066 GAGGAGCTGGTCCCTATATCCAGTGTGTTGCTTCCCAACACCCGAAACAAAGAAATCTC 1125
 QY 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394

Db 1126 GCCTTTGTGTGTCGCGCTAGCATGCCGCAACAGGCTATTAGTTGTGATCT 1185
 QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer 414
 Db 1186 TTGTCTGAAGCTCAAAATATGATCATCATCCAGAGACTACTAAGAGAGAGACTACC 1245
 QY 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
 Db 1246 AACAGATCAACAGT-----ATATTTCAAGACAAGCTTGGGATCT 1287
 QY 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIle 454
 Db 1288 TTATGGCCACCAAGAAAGAGAGAGAGATCTTCTCTTTTGGTCTTGCACATCATA 1347
 QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
 Db 1348 GTTCAATTCGATACCGAAGGATTAGAAGCTTCTCCGACTTCTCCGCCCTTCCAAA 1407
 QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
 Db 1408 TGGATGTGGCAAGGTTTCTTAGGATCAACATTAACATCAGGAGATCTCGTTCTTTGCT 1467
 QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
 Db 1468 TTATACATGTTCGTCATTTCCACCAACAATTTGAGAAAAGTCTCATCAATCATCTCATC 1527
 QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
 Db 1528 TCTGATCAACCGGAGCAACCATGATAAAACCTATCTCAAA 1569
 RESULT 4
 AY040024 1836 bp mRNA linear PLN 24-APR-2002
 LOCUS Arabidopsis thaliana putative lycopene epsilon cyclase (At5g57030)
 DEFINITION mRNA, complete cds.
 ACCESSION AY040024
 VERSION AY040024.1 GI:14532799
 KEYWORDS FLI-CDNA.
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Rosidiales; Magnoliophyta; eudicotyledons; core eudicots;
 Spermatophyta; Eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 1836)
 AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
 Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
 Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T.,
 Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
 Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
 Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
 Davis,R.W., Ecker,J.R. and Theologis,A.
 Arabidopsis Full Length cDNA Clones
 Unpublished
 2 (bases 1 to 1836)
 TITLE Submitted (11-JUN-2001) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 JOURNAL RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of AFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.
 Direct Submission
 TITLE Submitted (11-JUN-2001) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 JOURNAL RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of AFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S. X., Sakano, H., Pham, P. K., Bann, J., Chung, M. K., Dale, J. M., Gibson, H. A., Goldsmith, A. D., Jiang, P. X., Lee, J. M., Quach, H. L., Tang, C. C., Toriumi, M., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J., Mays, M. C., Miranda, M., Nguyen, M., Palm, C. J., Shinn, P. J., Southwick, A., Tracy, S. E., Davis, R. W., Ecker, J. R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

source

```
Location/Qualifiers
1. .1836
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
```

```

/clone="RAFL05-09-E07 (R10203)"
/note="This clone is in a modified pBluescript vector
(Lambda Zap) as a XhoI/SstI insert."

```

gene

```
ecotype: Columbia
1. .1836
/gene="At5g57030"
1. .105
```

1
4
•
8

```
1.2
/gene="At5g57030"
```

/note="compared to genomic sequence"

cps

```
106. .1680
/gene="At5g57030"
/codon_start=1
/evidence=experimental
/product="putative lycopersicon protein"
/protein_id="AAK64181.1"
/db_xref="GI:14532800"
```

[illegible]

3' UTR

1681. .1836
/gene="At5g57030"
324 c 440 g 542 t

324 C 440 g

Alignment Scores:

Pred. NO.:	3.le-159	Length:	1836
Score:	1966.50	Matches:	377
Percent Similarity:	81.84%	Conservative:	60
Best Local Similarity:	70.60%	Mismatches:	80
Query Match:	71.69%	Indels:	17
DB:	8	Gaps:	5

US-09-701-395A-23 (1-529) x AY040024 (1-1836)

QY 1 MetCluLeuLeuClyValArgAsnLeu-----IleSerSerCysProValTrp 16

Db
106 ATGGAGTGTGTGGGCTAGGAATTTCCGAGCAATGGCGGTTTCAACATTTCCCGTCATGG 165

17 ThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGly 36

Db 166 AGTTGTCGAGGAAATTTCCAGTGGTTAAGACATACAGCTATAGGAATATTCCTGTTTCCCT

2yCGGTTTCGGT ZZ

37 SerSerCysArgValAsppheg[nva]Arqa[aaagcgcvcv-----Ser-C- -

b

 ---TTGTG---	<div style="background-color: black; width: 8px; height: 8px;"></div> -----ACMCMCAACAAGTCCCGCCTGGTG serglyserarg 54
db 226	

AGGCGCAGAGCTAGCGGGCGGAAGTTCGGGTAGTGAG 270

SS SerSerValAlaArgLysGluGluGlyPheIleLysAlaGlyGly 74

[illegible]

Db 1351 AACAGATCAACAGT-----AATATTTCAAGACAAAGCTTTGGGATAC 1392
Qy 435 LeuTyrProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIle 454
Db 1393 TTATGGCCACCAAGAAAGAAAGACAGAGAGATCTTCTCTTTGGTCTTGCACATCA 1452
Qy 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
Db 1453 GTTCAATTCCGATACCAAGGCAATTAGAAGCTTCTCCGTAATTTCTTCCGCTTCCAAA 1512
Qy 475 TrpMetTyrTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
Db 1513 TGGATGTGGCAGGGTCTTAGGATCAACATTAACATCAGGAGATCTCGTCTCTTTGCT 1572
Qy 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
Db 1573 TTATACATGTTGCTCATTTACCAACAAATTTGAGAAAGGTCCTCATCATCATCTCATC 1632
Qy 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
Db 1633 TCTGATCCAAACGGAGCAACCATGATATAAAACCTATCTCAAA 1674
RESULT 5
AR003634
LOCUS AR003634 1860 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5744341.
ACCESSION AR003634
VERSION AR003634.1 GI:3964893
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1860)
AUTHORS Cunningham, F. X. Jr. and Sun, Z.
TITLE Genes of carotenoid biosynthesis and metabolism and a system for
screening for such genes
JOURNAL Patent: US 5744341-A 1 28-APR-1998;
FEATURES Location/Qualifiers
source 1..1860
BASE COUNT 531 a 327 c 448 g 554 t
ORIGIN
Alignment Scores:
Pred. No.: 1.26e-158 Length: 1860
Score: 1959.50 Matches: 376
Percent Similarity: 81.65% Conservative: 60
Best Local Similarity: 70.41% Mismatches: 81
Query Match: 71.44% Indels: 17
Gaps: 5
US-09-701-395a-23 (1-529) x AR003634 (1-1860)
Qy 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTrp 16
Db 109 ATGGAGTGTGTGGGGCTAGGAATTCGCAGCAATGGCGGTTTCAACATTTCCGTCATGG 168
Qy 17 ThrPheGlyThrArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGly 36
Db 169 AGTTGTCGAGGAATTTCCAGTGGTTAAGAGATACAGGTATAGGAATATTCGTTTCGT 228
Qy 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerArg 54
Db 229 ---TTGTGT-----AGTGTACAGCTACGGCGGGGGAAGTTCCCGTAGTGAG 273
Qy 55 SerSerValAlaTyrLysGluGlyPheValAspGluAspPheIleLysAlaGlyGly 74
Db 274 AGTTGTGAGCGGTGAGAGAAGATTTCGTCAGCAAGAAGATTTCGTAAGAGCTGGTGT 333
Qy 75 SerGluLeuLeuPheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLeu 94
Db 334 TCTGAGATCTTATTGTTCAATGTCAGCAGCAACAAAGATATGGATGAACAGCTTAAGCTT 393
Qy 95 AlaAspLysLeuProIleProPheGlyGluSerValMetAspLeuValValIleGly 114
Db 394 GTTGATAAGTGGCTCTCTATATCAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 453
Qy 115 CysGlyProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGly 134
Db 454 TGTGGTCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 513
Qy 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLys 154
Db 514 CTCATTTGGTCCAGATCTCTCTTTACTAACAAATACGGTGTGGGAGAGATGAATTCAT 573
Qy 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAsp 174
Db 574 GATCTTTGGGCTGCAAAATGTAATAGCATGTTTGGAGAGAGACTATTGTGTATCTCGAT 633
Qy 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHis 194
Db 634 GATGACAAGCCTAATACCATTTGGCGGTGTATGGAAGAGTTAGTCGACGTTTGTCTCAT 693
Qy 195 GluGluLeuLysArgCysValGluSerGlyValSerGlyValSerLysValGlu 214
Db 694 GAGGAGCTTTGAGGAGGTGTGTCAGTCAAGTCTCTGACGCAATTAACGTCATT 753
Qy 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle 234
Db 754 AGCATAAACAGAAAGCTTCTGATGGCTTAGACTTTGTTGCTGTGACGCAATTAACGTCATT 813
Qy 235 ProCysArgLeuAlaThrValAlaSerGlyAlaSerGlyLysLeuLeuGluTyrGlu 254
Db 814 CCTCGAGGCTTCCACTGTTGCTTCTGGAGCAGCTTCGGGAAGCTTTGCCAATACGAA 873
Qy 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsn 274
Db 874 GTTGGTGGACCTACAGATCTGTGTCAAACTGCATACGGCGTGGAGGTGGAGTGGAAAT 933
Qy 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLys 294
Db 934 AGTCCATATGATCCAGATCAATGTTTTCATGATTACAGAGATTATTAACGAGAAA 993
Qy 295 LeuGlnCysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThr 314
Db 994 GTTCGGAGCTTAGAAGCTGAGTATCAACGTTTCTGTCGCCATGCTTATGACAAAGTCA 1053
Qy 315 ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
Db 1054 AGACTCTCTCTCGAGGAGACATGTTGGCCCTCAAAAGATCTCATGCCCTTTGATTTCTA 1113
Qy 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
Db 1114 AAAACGAAGCTCATGTTAAGATTAGATACACTCGGAATTCGAATTTCAAAGACTTAGCAA 1173
Qy 355 GluGluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeu 374
Db 1174 GAGGAGTGGTCTTATATCCAGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1233
Qy 375 AlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394
Db 1234 GCCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1293
Qy 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer 414
Db 1294 TTGTCTGAAGCTCAAAATATGATCAGTCATCGCAGAGATACATAAGAGAAGAGACTACC 1353
Qy 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
Db 1354 AAACAGATCAACAGT-----AATATTTCAAGACAAAGCTTTGGGATAC 1395
Qy 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIle 454
Db 1396 TTATGGCCACCAAGAAAGAAAGACAGAGAGATCTTCTCTTTGGTCTTGCACATCA 1455
Qy 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474


```

QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
Db 1114 AAAACGAAGCTCATGTAAAGATTAGATACACGGAATTCGAATCTAAAGACTTACGAA 1173
QY 355 GluGluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGlnLysAsnLeu 374
Db 1174 GAGGAGTGTCTATATCCAGTTGGTGTCTCTGCCAACACCAACCAAGAAAGAAATCTC 1233
QY 375 AlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394
Db 1234 GCCTTTGGTGTCTCCGCTAGCATGTTACATCCGCAACAGGCTATTCAGTTGTGAGATCT 1293
QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer 414
Db 1294 TTGTCTGAAGCTCCAAAATATCATCAGTCATCGCAGAGATCTACTAGAGAAAGAGACTACC 1353
QY 415 AlaTyrValValSerGlyGlnSerAlaValAsnIleSerMetGlnAlaTyrSerSer 434
Db 1354 AAACAGATCAACAGT-----AATATTTCAGACCAAGCTTGGGATACT 1395
QY 435 LeuTrpProLysGluArgLysArgGlnAlaPhePheLeuPheGlyLeuGluLeuIle 454
Db 1396 TTATGCCACCAGAAAGAAAGACAGACAGCATCTCTCTCTTGTCTTGCACATCATA 1455
QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
Db 1456 GTTCAATTCATACCAAGGCAATAGAACGCTCTTCCGTACTTCTCCGCTTCCAAAA 1515
QY 475 TrpMetTrpGlyPheLeuGlySerSerLeuSerSerPheAspValLeuPheSer 494
Db 1516 TGGATGTGCAAGGGTTCTAGATCAACATTAACATCAGGAGATCTCCTCTCTTGTCT 1575
QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
Db 1576 TTATACATGTTGTCATTTCCACCAACAAATTTGAGAAAGGCTCATCATCATCTCATC 1635
QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
Db 1636 TCTGATCCAAACGGAGCAACCATGATATAAAACCTATCTCAA 1677

RESULT 7
AF321538 Locus AF321538 1780 bp mRNA linear PLN 28-FEB-2001
DEFINITION Lactuca sativa lycopene epsilon-cyclase mRNA, complete cds.
ACCESSION AF321538
VERSION AF321538.1 GI:12746312
KEYWORDS Lactuca sativa.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa.

REFERENCE
AUTHORS Cunningham, F.X. Jr. and Gantt, E.
TITLE One ring or two? Determination of ring number in carotenoids by lycopene varepsilon-cyclases
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (5), 2905-2910 (2001)
MEDLINE 21126984
PUBMED 11226339
REFERENCE
AUTHORS Cunningham, F.X. Jr.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2000) Cell Biology and Molecular Genetics, University of Maryland, Microbiology Building, Campus Drive, College Park, MD 20742, USA
FEATURES
source
location/Qualifiers
1. :1780
/organism="Lactuca sativa"
/cultivar="Romaine"
/db_xref="taxon:4236"
/clone="DY4"
/clone_lib="R.C. Bugos, H.Y. Yamamoto, Proc. Nat. Acad.
```

```

Sci. U.S.A. 93:6320-6325 (1996)"
77. :1678
/function="converts lycopene to epsilon-carotene via delta-carotene"
/notes="evidence by complementation in lycopene-accumulating Escherichia coli"
/codon_start=1
/evidence-experimental
/product="lycopene epsilon-cyclase"
/protein_id="AAK07434.1"
/db_xref="GI:12746313"
/translation="MECFGARNMTATMAVTCPRFTDCNIRHKFSLLKQRFTNLSAS
SLRQIKCSAKSDRCVYDKGIVSVADEEDYVKAGSELFEVQMRKSMSSQSLSEK
LAQIPGICILDLVITCGCPADLAALSAKGLNVLGIDPLPTFNNGVQDEFITG
LGLEGIEHSKDTLVYLDADADRIGRAYGRVHRLDHEELLRRRCVSGSVYLSKRV
ERTTEAPNGYSLIECEGNTIPCLRLATVAGSAGKFLVLELGGPRVCVOTAYGIEVE
VENNPDPDLMPMDYRDFSKHPESLEAKYPTFLYVMAKSPKIFFEETCLASREAM
PNNLKSMLSRKMGIRITRIEENSTIPVGGSLPNTQENLAFGAAASVHAFAP
GYSVRSLSSEAPNAAVIAKILRQDSKEMISLGYTNISKQAWETLWPLERKQRAF
FLFGLSHIVLMDLEGTFTFTFRLPQKMMWGLGSLSSLSLTDLIIFALYMFYIAPH
LRMELVRHLLSDPTGATMKAVLTI"
BASE COUNT 526 a 332 c 383 g 539 t
ORIGIN

Alignment Scores:
Pred. No.: 1-82e-156 Length: 1780
Score: 1934.00 Matches: 375
Percent Similarity: 78.69% Conservative: 57
Best Local Similarity: 68.31% Mismatches: 77
Query Match: 70.51% Indels: 40
DB: 8 Gaps: 5

US-09-701-395A-23 (1-529) x AF321538 (1-1780)
QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProVal 15
Db 77 ATGGAGTGTCTTGGAGCTCGAAACATGACGGCAACAATGGCGGTTTTTACGCGCCCTAGA 136
QY 16 TrpThr-----PheGlyThr 20
Db 137 TTCACGGACTGTAAATATCATCGCACAAATTTTCGTACTGAAACCAACCAAGATTACTAAT 196
QY 21 ArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerCysArg 40
Db 197 TTATCAGCATCGTCTTCGTTGCGTCAAAATTAAGTGCAGCGCTAAAGACGACCGTGTGTA 256
QY 41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLys 60
Db 257 GTGGAT-----AAA 265
QY 61 GluGlyPhe-----ValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeu 78
Db 266 CAAGGGATTTCCGTAGCAGACGAGAGAGATATGTGAGGCGCGGTGGATCGGAGCTGTTT 325
QY 79 PheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeu 98
Db 326 TTTGTTCATAATGCAGCGGACTAAGTCCATGGAAGCCAGCTAACTTCGGAAGAGCTA 385
QY 99 ProProIleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAla 118
Db 386 GCACAGATACCAATTCGAAATTCATGCTGTGTTGAGGCTGCTGCTGCTGCTGCTGCT 445
QY 119 GlyLeuSerLeuAlaAlaGluAlaAlaLysLeuGlyLysValGlyLeuIleGlyPro 138
Db 446 GGCTTGTCTTCTGCTGAGAGTGCAGCAAACTAGGTTGAGCTTGGACATTCATGGCCCT 505
QY 139 AspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeu 158
Db 506 GATCTTCTCTTTCAACAATTTATGGTGTGGCAGCATGAATTTATAGTCTTGGACTT 565
QY 159 GluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaPro 178
Db 566 GAAGGATGCAATTGAACATTTCTTGGAAAGATACCTCTTGTATACCTTGTATGCTGCTGCC 625
```

QY 179 ValLeuLeuGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeu 198
 Db 626 ATCCGCATAGTCTGTCATATGGCAGAGTTTCATCGTATCTACTTCAATGAAGAGTTGTTA 685
 QY 199 LysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGlu 218
 Db 686 AGAAGGTGTGGAATCAGGTGTTTCATATCTAAGTCTCAAGTAGAAGAAATCACTGAA 745
 QY 219 AlaGlyAspGlyHisSerLeuValCysGluAsnGluIlePheIleProCysArgLeu 238
 Db 746 GCTCCAAATGGCTATAGTCTCATGTAATGGAAGCAATATCACTATCCATCCAGGCTT 805
 QY 239 AlaThrValAlaSerGlyAlaSerGlyLysLeuGluTyrGluValGlyGlyPro 258
 Db 806 GCTACTGTTCATCAGGGCAGCTTCAGGCAATTTCTGGAGTATCACTTGGGGTCCC 865
 QY 259 ArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAsp 278
 Db 866 CGTGTGTGTCCTCAACACAGCTTATGATATAGAGTTGAGTTGAAACCAACCCCTATGAT 925
 QY 279 ProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSer 298
 Db 926 CCAGATCTAATGTGTTCATGATATAGAGCTTCTCAAAACATAAACCGGAATCTTTA 985
 QY 299 GluGluLysTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePhe 318
 Db 986 GAAGCAAAATATCCGACTTCTCTATGTCATGGCCATGCTCCAAACAAAATATCTTTC 1045
 QY 319 GluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLysArgLysLeu 338
 Db 1046 GAGGAACATGTTAGCTTCAAGAGAGCCATGCCCTTCAACTTCTCAAGTCCAAACATC 1105
 QY 339 MetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTyrPhe 358
 Db 1106 ATGTCACGATTAAGACGATGGTATCCGAATAACAGAACCTACGAGAGAAATGGTGC 1165
 QY 359 TyrIleProValGlyGlySerLeuProAsnThrGluGlnValThrLysAsnLeuAlaPheGlyAla 378
 Db 1166 TATATCCCGTAGTGGATCGTTACTTATACAGAACAAAGAAATCTCGCATTTGGTGC 1225
 QY 379 AlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAla 398
 Db 1226 GCAGCTAGTATGGTGCACCCCTGCCACAGGGTATTCAGTTGTCGATCTTGTCAAGACT 1285
 QY 399 ProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValVal 418
 Db 1286 CCTAATATGACGACGATCTTCTGAAGATTTTAAAGACAAATCAATCTAAAGAGATGAT 1345
 QY 419 SerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTyrSerSerLeuTyrProLys 438
 Db 1346 TCT---CTTGGAATAACACTAACATTTCAAAACAGCATCGGAAACATTTGGCCACTT 1402
 QY 439 GluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAsp 458
 Db 1403 GAAAGGAAACACACGAGCGCTTCTTCTATTCGAGCTATCACACATCGTCTAATGGAT 1462
 QY 459 IleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrp 478
 Db 1463 CTAGAGGGAACACGATCATTTTCCGACTTCTTCTGTTGCGCAATGGATGGTGG 1522
 QY 479 GlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPhe 498
 Db 1523 -GGATTTTGGGTCTCTCTTATCTTCAACGAGATTGTGATATATTTGGGCTTTATATGTTT 1582
 QY 499 ValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSer 518
 Db 1583 GTGATAGCACTCAGCTTGAAGTGAAGTGGTGTAGATCTTCTTCTGATCCGACA 1642
 QY 519 GlyAlaValMetValArgAlaTyrLeu 527
 Db 1643 GGGCACTATGTTAAAGCATATCTC 1669

RESULT 8
 AF251016
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AF251016 1830 bp mRNA linear PLN 19-NOV-2001
 Tagetes erecta epsilon cyclase mRNA, complete cds.
 AF251016
 AF251016.1 GI:9971815

Tagetes erecta.

ORGANISM

Tagetes erecta
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 Helianthaceae; Tagetes.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 (bases 1 to 1830)
 Moehs,C.P., Tian,L., Osteryoung,K.W. and DellaPenna,D.
 Analysis of carotenoid biosynthetic gene expression during marigold
 petal development
 Plant Mol. Biol. 45 (3), 281-293 (2001)
 21187422
 11292074
 2 (bases 1 to 1830)
 Moehs,C.P., Tian,L. and DellaPenna,D.
 Direct Submission
 Submitted (03-APR-2000) Biochemistry, Univ.of Nevada, Reno, NV
 89557, USA

FEATURES

Location/Qualifiers
 1..1830
 /organism="Tagetes erecta"
 /db_xref="taxon:13708"
 141..1691
 /codon_start=1
 /product="epsilon cyclase"
 /protein_id="AAG10438.1"
 /db_xref="GI:9971816"
 /translation="MSMRAGHTMATMAAFTCPRMTSIRYTKIKNAKSQLVVKE
 IEEDYVYKAGGSELLEFVOMQNKSDAQSSLOKLPVPIGGGSDNCILDLVIGC
 GPAGLALAGESAKILGNVALIGDPLPTNNYGVWDEFGIGLEGGCTEHWRTVVYL
 DDNDPILIGRAYGRVSRDLLEHLLTRCMESGVYLSKVERITEAPNGSLIECGN
 ITTCRLATVAGSAAGSKLLQYELGGPRVCVOTAYGIEVESIPDPSPLMFMDYRD
 YTKHSQSLAOPYTELYVMPSPTKVFFETCLASKEAMPFELKTKLSRLKTMGI
 RLNTKEEYSYIPVGGSLNTEQKNAFGAAAMVHPATGYSVRSLSAPNAAVI
 AKLIGKNSQMLDHDHRYTTNISKQAWETLWPLERKQRAFFLGLALIVQMDIEGTR
 TFFRTFRLPTMMWNGFLSGSLSDLIIFAFYFIIAPHSLRMGLVRHLLSDPTGTT
 MLKAYLTI"

BASE COUNT 537 a 351 c 422 g 520 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.89e-156 Length: 1830
 Score: 1934.00 Matches: 365
 Percent Similarity: 83.95% Conservative: 64
 Best Local Similarity: 71.43% Mismatches: 74
 Query Match: 70.51% Indels: 8
 DB: 8 Gaps: 2

US-09-701-395A-23 (1-529) x AF251016 (1-1830)

QY 22 AsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArgVal 41
 Db 159 CACATGACGCAACAATGGCGCTTTTACATGCCCTAGGTTTATGATCATGCATCATC 218
 QY 42 AspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLysGlu 61
 Db 219 ACGAAGCAAAATTAAGTGC-----AAGCGTCTAAAGCCAGCTAGCTGTTAAACAA 269
 QY 62 GlyPheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuPheValGln 81
 Db 270 GAGATTGAGGAGGAAGAGATTATGTGAAAGCGGTGGATCGGAGCTGCTTTTGTCAA 329
 QY 82 MetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProIle 101
 Db 330 ATGCAACAGAAATAAGTCCATGGATGCACAGCTAGCTATCCCAAAAGCTCCCAAGGGA 389

QY 102 PropheGly-----GluSerValMetAspLeuValValIleGlyCysGly 116
Db 390 CCAATAGGAGGAGGAGGAGACAGTAACCTGTATACATGATTTGGTTGTAATGCTTGGT 449
QY 117 ProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuValGlyLeuIle 136
Db 450 CCGTGTGGCTTGGCTTCTGCTGAGAAATCAGCAAGCTAGGCTGAATGTCGCACTATC 509
QY 137 GlyProAspLeuProPheThrAsnAsnTyrGlyValTyrGluAspGluPheLysAspLeu 156
Db 510 GGCCTGTATCTCCCTTTTCAAAATACTATGTTGGAGGAGTAATTTATAGTCTT 569
QY 157 GlyLeuGluArgCysIleGluHisAlaTyrLysAspThrIleValTyrLeuAspAsnAsp 176
Db 570 GGACTTGGGGCTGTATTCAGCATGTTTGGCGAGATCTGTAGTATATCTTGTATGACAC 629
QY 177 AlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGlu 196
Db 630 GATCCCATTTCTCATAGGTCGTGCTATGACGAGTTAGTCGTGATTTACTTTCACGAGGAG 689
QY 197 LeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIle 216
Db 690 TTGTTGACTAGTGCATGAGTGCAGCGCTTCATATCTAGCTCCAAAGTGGACCGATT 749
QY 217 ThrGluAlaGlyAspCysHisSerLeuValValCysGluAsnGluIlePheIleProCys 236
Db 750 ACTGAAGCTCCAAATGGCCTAAGTCTCATAGAGTGTGAAGGCAATATACAAATTCATGC 809
QY 237 ArgLeuAlaThrValAlaSerGlyAlaLysSerGlyLysLeuLeuGluTyrGluValGly 256
Db 810 AGGCTTGCTACTGTCGCTTCTGAGCAGCTTCTGGAACCTTTTGCAGTATGAATTCGC 869
QY 257 GlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnPro 276
Db 870 GTCCCGCTGTTGGCTTCAACAGCTTATGTTATAGAGTTGAGTTGAAAGCATACCC 929
QY 277 TyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnLysLeuGln 296
Db 930 TATGATCCAAAGCTAATGCTTTCATGGATTATAGAGCTACACCAACATAAATCTCAA 989
QY 297 CysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeu 316
Db 990 TCACCTAGAAGCACAAATATCCAACTTTTGTATGTCATGCCAATGCTCCAACTAAAGTA 1049
QY 317 PhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArg 336
Db 1050 TTCCTTGAGAACTGTTTGGCTTCAAAAGAGCCATGCCCTTTCAGTTATTAAGACA 1109
QY 337 LysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGlu 356
Db 1110 AAACCTCATGCTCAAGATTAAAGACTATGGGGATCCGAATAACCAAACTTATGAAGAGAA 1169
QY 357 TrpSerTyrIleProValGlyGlySerLeuProAsnThrGluClnLysAsnLeuAlaPhe 376
Db 1170 TGGTCATATATTCAGTAGTGGATCCTTACCATAATACCAGCAAGAACCTTGTGATTT 1229
QY 377 GlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSer 396
Db 1230 GGTGTGCTGTAGCATGTTGATCCAGCCACAGGATATCGGTTGTAGATCAGCTCA 1289
QY 397 GluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyr 416
Db 1290 GAAGCTCCTAATATGTCAGCAGTAATTCGAAAGATTTTGGGAAAGGAAATTCAAAACAG 1349
QY 417 ValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrp 436
Db 1350 ATGCTGTGATGGAAGATACCAACCACTCTCAAGCAAGCTTGGGAAACACTTGG 1409
QY 437 ProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGln 456
Db 1410 CCCCTTGAAGGAAGACAGAGAGCATCTTCTCTTTGGATTACACACTGATGTCCAG 1469
QY 457 LeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMet 476

Db 1470 ATGATATTAGGGGACCCGACACATCTTCCGACATTTCTCCGCTTCCACATGATG 1529
QY 477 TrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyr 496
Db 1530 TGGTGGGGTTCCTTGGATCTTCGTTATCATCAACTGACTTGATAAATATTTGCGTTTAC 1589
QY 497 MetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAsp 516
Db 1590 ATGTTTATCATAGCACCGCATAGCTGAGAATGGGTCTGGTTAGACATTTGCTTTCTGAC 1649
QY 517 ProSerGlyAlaValMetValArgAlaTyrLeu 527
Db 1650 CCGACAGAGGAACAATGTTAAAGCGTATCTC 1682
RESULT 9
ARI52414 ARI52414 1887 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 5 from patent US 6232530.
DEFINITION ARI52414
ACCESSION ARI52414
VERSION ARI52414.1 GI:15118464
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1887)
AUTHORS Dellapenna, D. and Cunningham, F. X. Jr.
TITLE Marigold DNA encoding beta-cyclase
JOURNAL Patent: US 6232530-A 5 15-MAY-2001;
FEATURES Location/Qualifiers
source 1. .1887
BASE COUNT 561 a 366 c 428 g 532 t
ORIGIN
Alignment Scores:
Pred. No.: 1.97e-156 Length: 1887
Score: 1934.00 Matches: 365
Percent Similarity: 83.95% Conservative: 64
Best Local Similarity: 71.43% Mismatches: 74
Query Match: 70.51% Indels: 8
DB: 6 Gaps: 2
US-09-701-395a-23 (1-529) x ARI52414 (1-1887)
QY 22 AsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArgVal 41
Db 159 CACATGACGGCAACAATGGCGCTTTTACATGCGCTTAGTTTATGACTAGCATCAGATAC 218
QY 42 AspPheGlnValArgAlaAspGlyGlySerArgSerSerValAlaTyrLysGlu 61
Db 219 ACGAAGCAATTAAGTGC-----AACGCTGCTAAAGCCAGCTAGTCTGTAAACAA 269
QY 62 GlyPheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheValGln 81
Db 270 GAGATTGAGGAGGAAGAAGATTATGTAAAGCGGTGGATCGGAGCTGCTTTTGTCAA 329
QY 82 MetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPhe 101
Db 330 ATCAACAGAAATAGTCCATGGATGTCAGCTAGCTATCCCAAGAGCTCCCAAGGTA 389
QY 102 ProPheGly-----GluSerValMetAspLeuValIleGlyCysGly 116
Db 390 CCAATAGCAGGAGGAGGAGGACAGTAACCTGTATCTGATGTTGGTTGTTGGT 449
QY 117 ProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGlyLeuIle 136
Db 450 CCGTGTGGCTTGGCTTCTGCTGAGAAATCAGCAAGCTAGGCTTGAATGTCGCACTATC 509
QY 137 GlyProAspLeuProPheThrAsnAsnTyrGlyValTyrGluAspGluPheLysAspLeu 156
Db 510 GGCCTGTATCTCCTTTTACAAATAACTATGTTGGGAGGATGAATTTATAGTCTT 569

```

QY 157 GlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAsp 176
Db 570 GGACTTGGAGGCTCTATTGAACATGTTGGCGAGATCTAGTATATCTTGATGACAC 629

QY 177 AlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHisGluGlu 196
Db 630 GATCCCATCTCTATAGGTCGTGCTATGACGAGTGTAGTCGTGATTTACTTCCAGGAG 689

QY 197 LeuLeuLysArgCysValGluSerGlyValSerTyrLysSerLysValGluArgIle 216
Db 690 TTGTTGACTAGTGCATGAGTCAGCGCTTCATATCTGAGTCCAAAGTGAACGGATT 749

QY 217 ThrGluAlaGlyAspGlyHisSerLeuValCysGluAsnGluIlePheIleProCys 236
Db 750 ACTGAAGCTCCAAATGGCTTAAGCTCATAGAGTGTGAAGCAATATCAATTCATCC 809

QY 237 ArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGly 256
Db 810 AGGCTTGCTACTGCTCTGAGCAGCTTCTGGAAACTTTTCAGTATGAATCTTGGC 869

QY 257 GlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnPro 276
Db 870 GGTCCCGGTTGTTGGTTCAAACAGCTTATGCTATAGAGTTGAGTTGAAGCATACC 929

QY 277 TyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGln 296
Db 930 TATGATCCAAAGCTTAATGGTTTCATGATATAGAGTATACCAACATATAATCTCAA 989

QY 297 CysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeu 316
Db 990 TCACCTAGAAGACATATCCAACTTTTGTATGTCATGCCAATCTCTCCAACTAAAGTA 1049

QY 317 PhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLysArg 336
Db 1050 TTCCTTGAGGAACCTGTTGGCTTCAAAGAGGCCATCGCTTTGAGTTATTGAAGACA 1109

QY 337 LysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGlu 356
Db 1110 AAACCTCATGCAAGATTAAAGCATATGGGATCCGAATACCAAACTATGAAGAGAA 1169

QY 357 TrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPhe 376
Db 1170 TGGTCATATATCCAGTAGTGGATCTCTACCAATACCGAGCAAAAGAACCTTGATTT 1229

QY 377 GlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSer 396
Db 1230 GGTGCTGCTGCTAGCATGCTCCAGCCAGCATATCGGTTGTAAGATCATCTCTCA 1289

QY 397 GluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyr 416
Db 1290 GAAGCTCCTAATATTGACAGCAGTAATTCCAAAGATTTTAGGAAGGAATTCAAACAG 1349

QY 417 ValValSerGlyGlnSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrp 436
Db 1350 ATGCTTGATCATGGAAGATACACAACTCTCAAGCAAGCTTGGAAACACTTTGG 1409

QY 437 ProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuValGln 456
Db 1410 CCCCTTGAAGGAAACACAGACAGACATCTTCTCTTTGGATTGACATGATGTCACG 1469

QY 457 LeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMet 476
Db 1470 ATGGATATTGAGGGACCCGACATCTTCCGACCTTCTCCGCTTGGCCACATGGATG 1529

QY 477 TrpTrpGlyPheLeuGlySerLeuSerSerPheAspLeuValLeuPheSerMetTyr 496
Db 1530 TGGTGGGGCTTCTTGGATCTTCTTATCATCACTACATGATGATATATTGCGTTTAC 1589

QY 497 MetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisIleLeuSerAsp 516
Db 1590 ATGTTTATCATAGCAGCGCATAGCCTGAGAAATGGGCTGTTAGACATTTGCTTTCTGAC 1649

QY 517 ProSerGlyAlaValMetValArgAlaTyrLeu 527

```

Db 1650 CCGACAGGAGGACAAATGTTAAAGCGTATCTC 1682
 RESULT 10
 AY099485
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Tagetes erecta
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Tagetes.
 1 (bases 1 to 1916)
 Del Villar-Martinez,A.A., Garcia-Saucedo,P.A., Cruz-Hernandez,A.
 and Paredes-Lopez,O.
 Isolation and functional characterization of carotenogenic genes
 (psyl, pds, lcy-b and lcy-e) from marigold (Tagetes erecta) and
 ultrastructural changes during flowering development
 Unpublished
 2 (bases 1 to 1916)
 Del Villar-Martinez,A.A., Garcia-Saucedo,P.A., Cruz-Hernandez,A.
 and Paredes-Lopez,O.
 Direct Submission
 Submitted (24-APR-2002) Bioquimica y Biotecnologia, Cinvestav
 Unidad Irapuato, Km. 9.6 Lib. Nte. Carr. Irapuato-Leon, Irapuato,
 Guanajuato, Mexico
 Location/Qualifiers
 1..1916
 /organism="Tagetes erecta"
 /db_xref="taxon:13708"
 1..1916
 /gene="lcy-e"
 173..1750
 /gene="lcy-e"
 /note="lcy-e"
 /codon_start=1
 /product="epsilon cyclase"
 /protein_id="AA045382.1"
 /db_xref="GI:21360359"
 /translation="MSMRAGHTMATMAAFTCPREMTSIRYTKIOKCNAAKSQLVVKE
 IEEDYVRKAGSELLFVMOONKSDAQSSLOKLPVPVIGGSDSNCLDLVWIGC
 GPAGLALAGESAKLGNVALIGDLPFTNNYGVWDEFGLEGCGIEHWRTVYVL
 DNDPTILIGRAYGRVSRDLLEHLLTCHESGVYSSKVERITEAPNGULIECEGN
 ITPCRLATVASGAASGKLLQYELGGPRVCVQYAYGYEVEISIPYDPSLMVPMYRD
 YTKHSQLEAQPTFLYVMPSPSTKVFVEETCLASKEAMPFELLTKMSRLKTMGI
 RTKTYEYLVAQYLEWSPYIPVGSLEPTEOKNLAFGAAASNVHPATGYSVVRLS
 EAPNYAAVIAKILGKNSQMLDLGRYTNISKOAWETLWPLERKQRAFFLFLGLALI
 VOMDIEGTRTFRTFFPLPTMMWGLGSSLSSTDLIIFAFYFIITAPHSLRMLGVRH
 LLSDPDTGGTMLKAYLTI"
 BASE COUNT 580 a 361 c 427 g 548 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1..15e-154 Length: 1916
 Score: 1913.50 Matches: 365
 Percent Similarity: 82.31% Conservative: 63
 Best Local Similarity: 70.19% Mismatches: 75
 Query Match: 69.76% Indels: 17
 DB: 8 Gaps: 3
 US-09-701-395A-23 (1-529) x AY099485 (1-1916)
 QY 22 AsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArgVal 41
 Db 191 CACATGACGCAACATGCGCGCTTTTACATGCCCTTAGGTTTATGATGATCAGATC 250
 QY 42 AspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLysGlu 61
 Db 251 ACGAAGCAAAATTAAGTGC-----AACGCTGTAAAGCCAGCTAGTCGTTAAACAA 301

Qy	62	GlypheValAspGluGluAspPheIleTysAlaGlyGlySerGluLeuLeuPheValGln	81
Db	302	GAGATTGAGGAGGAAGAAGATTATGTGAAGCCGGTGGATGGAGCTGCTTTGTTCAA	361
Qy	82	MetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPheIle	101
Db	362	ATGCAACAGAAATAGTCCATGGATGCACAGTCACTCCCTATCCCAAAAGCTCCCAAGGTA	421
Qy	102	ProPheGly-----GluSerValMetAspLeuValIleGlyCysGly	116
Db	422	CCAAATAGGAGGAGGAGACAGTAACTGTATACTGCATTTGGTTGTAATTTGGTTGGT	481
Qy	117	ProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGlyLeuIle	136
Db	482	CTGTCTGGCCCTGCTCTGCTGGAGAATCAGCCAAGCTAGGCTTGAATGTCCGCACTTATC	541
Qy	137	GlyProAspLeuProPheThrAsnAsnTyrGlyValTTPGluAspGluPheLysAspLeu	156
Db	542	GGCCCTGATCTCTTTACAATAACTATGGTGTGGGAGGATGAATTTATAGTCTT	601
Qy	157	GlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAsp	176
Db	602	GGACTTTGAGGGCTGTATTGAACATGTTGGCGAGATACTGTAGTATATCTTTGATGACAAC	661
Qy	177	AlaProValIleLeuGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGlu	196
Db	662	GATCCCAATCTCATAGGTGGTCCATGGACGAGTTAGTCGTGATTACTTTCACGAGGAG	721
Qy	197	LeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIle	216
Db	722	TTGTTGACTAGGTGCTGAGTCAGCGCTTCATATCTGAGCTCCAAGTGAACGCGATT	781
Qy	217	ThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCys	236
Db	782	ACTGAAGCTCCAAATGGCTTAAGTCTCATAGAGTGTGAAGGCAATATCACAAATTTCCATGC	841
Qy	237	ArgLeuAlaThrValAlaSerGlyAlaLaserGlyLysLeuLeuGluTyrGluValGly	256
Db	842	AGGCTTGCTACTGCTGCTCTGGAGCAGCTTCGGGAACCTTTGCAGTATGAACCTTTGGC	901
Qy	257	GlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnPro	276
Db	902	GGTCCCGCTGTTTGGCTTCAACAGCTTATGGTTACGAGGTGAGGTGTAAGAGCATACCC	961
Qy	277	TyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnLysLeuGln	296
Db	962	TATGATCCAAAGCCTAATGTTTTTCATGGATTATAGAGACTACACCAATAATATCTCAAA	1021
Qy	297	CysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeu	316
Db	1022	TCACTAGACACAAATATCCACATTTTGTATGTCATGCCAATGTCTCCAACTPAAAGTA	1081
Qy	317	PhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArg	336
Db	1082	TTCTTTGAGAAACTTGTTTGGCTTCAAAAGAGGCCATGCTCTTTGAGTTATTCAAGACA	1141
Qy	337	LysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu-----	354
Db	1142	AAACTCATGTCAAGATPAAAGACTATGGGGATCCGAATTAACCAAACTTTATGAAGAGTAT	1201
Qy	355	-----GluGluTrpSerTyrIleProValGlyGlySerLeuPro	367
Db	1202	CTTGTGCTGTCAATATTGGAGAAATGGTCATATATTCAGTAGTGGTGGATCCCTTCCA	1261
Qy	368	AsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaLaserMetValHisProAlaThr	387
Db	1262	AATACCGAGCAAAAGAACCTTGCATTTGGTGTGCTGTAGCATGGTGTGCATCCAGCCACA	1321
Qy	388	GlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLys	407
Db	1322	GGATATTCGGTGTGAAGTACTGTGAGAAGCTCTTAATATGACGAGTAATGTGCAAG	1381

Qy	408	IleLeuIysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAlaIle	427
Db	1382	ATTTTAGGAAAGAAATTTTAAACACAGATGCTTGTGGAAGATACACAAACACATC	1441
Qy	428	SerMetGlnAlaTyrSerSerLeuTyrProLysArgLysArgGlnArgAlaPhePhe	447
Db	1442	TCAAAGCAAGCTTGGGAACACTTTGGCCCTTGAAGGAAAGACAGAGAGCATTTCTTT	1501
Qy	448	LeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhePheArg	467
Db	1502	CTCTTTGGATTAGCACTGATTGTCAGATGATATTGAGGGACCGCACATTTCTCCGG	1561
Qy	468	ThrPheArgLeuProThrTyrMetTyrTrpGlyPheLeuGlySerSerLeuSerSer	487
Db	1562	ACTTCTTCGCCCTGGCCACATGATGTGTGGGGTTTCTTGGATCTTCGTTATCATCA	1621
Qy	488	PheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMet	507
Db	1622	ACTGACTGTGATAATTTTGCCTTTTACATGTTTATCATAGCAGCGCATGCTGAGAATG	1681
Qy	508	SerLeuValArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeu	527
Db	1682	GGTCTGGTTAGACATTTGCTTCTTGACCGCAGAGGAGACAATGTTAAAGCGTATCTC	1741
RESULT	11		
LOCUS	LELYCOCYC	1697 bp	mRNA linear
DEFINITION	Lycopersicon esculentum	mRNA for lycopene epsilon-cyclase.	
ACCESSION	Y14387		
VERSION	Y14387.1	GI:3005982	
KEYWORDS	CrtL-e-1 gene; lycopene epsilon-cyclase.		
SOURCE	Lycopersicon esculentum.		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.		
REFERENCE	1 (bases 1 to 1697)		
AUTHORS	Ronen, G., Cohen, M., Zamir, D. and Hirschberg, J.		
TITLE	Regulation of expression of the gene for lycopene epsilon cyclase during fruit ripening of tomato		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1697)		
AUTHORS	Hirschberg, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JUL-1997) J. Hirschberg, The Hebrew University of Jerusalem, Dept of Genetics, Givat Ram, Jerusalem 91904, ISRAEL		
FEATURES	Location/Qualifiers		
source	1..1697		
	/organism="Lycopersicon esculentum"		
	/cultivar="VF36"		
	/db_xref="taxon:4081"		
	/tissue_type="leaf"		
	6..1586		
gene	/gene="CrtL-e-1"		
CDS	6..1586		
	/gene="CrtL-e-1"		
	/function="converts lycopene to delta-carotene"		
	/codon_start=1		
	/product="lycopene epsilon-cyclase"		
	/protein_id="CAA74745.1"		
	/db_xref="GI:3005983"		
	/db_xref="SPTREMBL:O65937"		
	/translation="MECVGVQNVGMAYLTPRLNRWSGELCOEKSIFLAYBOYESKQSSSDSCVDKDFADDEYIKAGSQLVQMDQKMDQKSLDSLELRISAGCTTDLIVDTGSPAGLAAESAKLGNVLGVPDLPFTNNYGVWDEFKDLGLQACI EHWYRDTIVYLDDEPDLIAGRYSRHFLHEELLRRCVEAGLVYLNKSKVDRIINP NGOSLVECDGVEYIPCRFTVYSAASGKFLQYELGSPRVSVOTAYGVDEVNDNEP PSLWPMVDYLRHDAQSLKALYFELYAMKMSPTVFEECTLASKDAMPEDLLK KLMLNLTGVRILKEIYEEWSYIPVGGSLPTEQKTLAFGAASVWHPATGYSVYRS LSBAPKCAVNLILRHYSKNMLTSSSIPSTQAWNTLPOEKRQRQSFFLGLAL ILQDLGIRSFYRFRFVPEKMQWQGLGSSLSADMLFAFTYMEITAPNDMRKGLIR HILSDPTGFLIRTYLTF"		

BASE COUNT 473 a 308 c 403 g 513 t

ORIGIN

Alignment Scores:

Pred. No.: 2.16e-154 Length: 1697

Score: 1909.50 Matches: 368

Percent Similarity: 81.19% Conservativeness: 68

Best Local Similarity: 68.53% Mismatches: 78

Query Match: 69.61% Indels: 23

DB: 8 Gaps: 6

US-09-701-395a-23 (1-529) x LELYCOCYC (1-1697)

QY 1 MetGluLeuLeuGluValAlaArgAsnLeuLeuSerSerCysProVal-----15

DB 6 ATGGAGTGTGTGGAGTTGCAAAATGTTGGAGCAATGCGACTTTTAACCGCTCCGAGATTG 65

QY 16 -----TrpThrPheGlyThrArgAsnLeuSerSerLys-----LeuAlaTyr 30

DB 66 AACCGTTGGTGGGAGGA-----GAGTTATGCCAAGAAAAGCATCTTTTGGCGGTAT 119

QY 31 AsnIleHisArgTyrGlySerSerCysArgValAspPheGlnValArgAlaAspGlyGly 50

DB 120 -----GAGCAGTATGAAGTAATGT-----AATGACAGT 149

QY 51 SerGlySerArgSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIle 70

DB 150 AGTGTAGTACAGCTGTGTAGTTGATGAAGAGATTTGCTGATGAAGAGATTATATA 209

QY 71 LysAlaGlyGlySerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLys 90

DB 210 AAAGCCGGTGGTTCGCAACTGTATTGTTGTTCAATGCAGCAGAAAAGATGATGATCAG 269

QY 91 GlnAlaLysLeuAlaAspLysLeuProPheGlyGluSerValMetAspLeu 110

DB 270 CAGTCAAGCTTCTGATGAGTTACAGCAAAATCTGCTGGCAGCAACCTGACTGATTA 329

QY 111 ValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaGluAlaLysLeuGly 130

DB 330 GTGGTAATCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389

QY 131 LeuLysValGlyLeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGlu 150

DB 390 TTGAAGTGGGCTGTTGGGCTGATCTTCCTTTTCAAAACAACATGATGGTGGAG 449

QY 151 AspGluPheLysAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIle 170

DB 450 GACGAGTTCAAAAGATCTTGGTCTTCAAGCCTGCATGAACTGTTGGCGGGATACCAT 509

QY 171 ValTyrLeuAspAsnAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArg 190

DB 510 GTATATCTGTATGATGATGAACCTATTCTTATGGCGCTGCTATGGAAGATTAGTCGC 569

QY 191 HisLeuLeuHisGluGluLeuLysArgCysValGluSerGlyValSerTyrLeuAsp 210

DB 570 CATTTTCTGCAGGAGGATTTACTCAAAAGGTGTGTGGAGCAGGTGTTGTATCTAAAC 629

QY 211 SerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsn 230

DB 630 TCGAAAGTGGATGATGATTTGGGCCACAAATGGCCAGAGTCTTGTAGAGTGGAGGT 689

QY 231 GluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeu 250

DB 690 GATGTTGATTCCTCCAGGTTTGTGACITGTGTCATGGGGGAGCCTCGGGGAAATTC 749

QY 251 LeuGluTyrGluValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluVal 270

DB 750 TTGCAGTATGAGTTGGAGTCTAGAGTTCTGTTCAACACCTTATGGAGTGGAGTT 809

QY 271 GluValGluAsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyr 290

DB 810 GAGGTTGATACAAATCCATTGACCCGAGCCTGATGTTTTCATGATATAGATTAT 869

QY 291 MetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetPro 310

DB 870 CTCAGACACGACGCTCAATCTTTAGAGCTAAATATATCAACATTTCTTTATGCGATGCC 929

QY 311 MetSerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetPro 330

DB 930 ATGCTCCACACAGAGCTTTTTCGAGGAACTTTGTTGGCTTCAAAAGATGCAATGCCA 989

QY 331 PheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThr 350

DB 990 TTCGATCTGTAAAGAAAAAAGCTGATGCTGATGCAACCCCTGGTGTAGAATAAA 1049

QY 351 LysValTyrGluGluGluThrPheTyrIleProValGlyGlySerLeuProAsnThrGlu 370

DB 1050 GAAATTTACGAGGAGGAATGGCTTACATACCGGTTGGTGGATCTTTGCCAAATACAGAA 1109

QY 371 GlnLysAsnLeuAlaPheGlyAlaAlaLaSerMetValHisProAlaThrGlyTyrSer 390

DB 1110 CAAAAACACTTTCATTTGCTGCTGCTAGCATGGTTTCATCCAGCCACAGGTTATTCA 1169

QY 391 ValValArgSerSerSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLys 410

DB 1170 GTGCTGAGATCACTTCTGAAGCTCCAAATCGCCTCTGTACTTGCAAAATATATAGA 1229

QY 411 GlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGln 430

DB 1230 CAACATTATAGCAAGACATGCTTACC---AGTTCAAGTATCCCGAGTATATCAACTCA 1286

QY 431 AlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGly 450

DB 1287 GCTTGAACACTCTTTGGCCACAAGAACGAAACGACAAAGATCGTTTTCTCTATTGGA 1346

QY 451 LeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePhe 470

DB 1347 CTGCTCTGATATTGCGAGCTGGATTTGAGGGGATAGAGTCATTTTCCGGCATTTCTC 1406

QY 471 ArgLeuProThrTrpMetTrpTyrGlyPheLeuGlySerSerLeuSerSerPheAspLeu 490

DB 1407 CGTGTGCAAAATGGATGGCAGGATTTCTTGGTTCAAGTCTTTCTTCAGCAGACCTC 1466

QY 491 ValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeu 510

DB 1467 ATGTTATTTGCTTCTACATGTTTATTATTGACCAATGACATGACAAAGGCTTGATC 1526

QY 511 ArgHisLeuLeuSerAspProSerGlyValAlaValMetValArgAlaTyrLeu 527

DB 1527 AGACATCTTTTATCTGATCTACTGCTGCAACATTGATAAGAACTTATCTT 1577

RESULT 12

AF463497

LOCUS

DEFINITION

cgs.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 2357

/organism="Spinacia oleracea"

Spinacia oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Caryophyllaceae; Chenopodiaceae; Spinacia.

1 (bases 1 to 2357)

DeSouza, M.L., Kollmann, S.R. and Schroeder, W.A.

Production of Lutetin in Microorganisms

Unpublished

2 (bases 1 to 2357)

DeSouza, M.L., Kollmann, S.R. and Schroeder, W.A.

Submitted (27-DEC-2001) Central Research, Cargill Inc., P.O. Box 5702, Minneapolis, MN 55440, USA

Location/Qualifiers

1. 2357

/organism="Spinacia oleracea"

Rosidae; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 1550)
Costa, M.C., Otoni, W.C. and Moore, G.A.
Direct Submission
Submitted (22-FEB-2002) Horticultural Sciences, University of
Florida, 1301 Fifield Hall, Gainesville, FL 32611, USA
FEATURES
Location/Qualifiers
1..1550
/organism="Citrus x paradisi"
/cultivar="Duncan"
/db_xref="taxon:37656"
45..1358
/codon_start=1
/product="lycopen epsilon-cyclase"
/protein_id="AAU92114.1"
/db_xref="GI:19569601"
/translation="MDMFLPISIGNGILDVVGCGPAGLAAAEAKLGNLVGIL
GPDLPNNYGVDEDFLEGECIEHWVRDVTWYIDEPIILIGRAYGRVSRHLIL
ELLRCFYCVESGVLSKSEVITESTSHRLVACEHDMIVPCRLATVASGAASKLLE
YEVGPKSVQTVAYGVEVENNNPDPLSMVENDYDCTQEVPSFESDNPFLIYVMP
NSSTVFVEETCLASKDGLRDLILKKLMARLERGLQVLTTEEWSYIPVGGSLPN
TSQNALPFAAASVMVHQATGYSVVSSEAPNTASATILKHHDSRGRLTHEQSNEN
LEMAWNTLWPOEKQRQRAFFLFGALILQDLQIEGTFRTFLPFLKMMWHGFLGSS
LSSADLILFIMFLIAPNDLNRKLIHLVSDPTGATMVTYTLT"
BASE COUNT 469 a 270 c 333 g 478 t
ORIGIN
Alignment Scores:
Pred. No.: 7,49e-144 Length: 1550
Score: 1786.00 Matches: 333
Percent Similarity: 87.6% Conservative: 44
Best Local Similarity: 77.44 Mismatches: 50
Query Match: 65.1% Indels: 3
DB: 8 Gaps: 0
US-09-701-395A-23 (1-529) x AF486650 (1-1550)
QY 98 LeuProProlleProPheGlyGluSerValMetAspLeuValValleGlyCysGlyPro 117
Db 60 CTACCGCCAAATACATTTGGTAATGGTATTTTGGATTTGGTGTATGGTTGTGGCCCA 119
QY 118 LAAGlyLeuSerLeuAlaLagLuAlaLysLeuGlyLeuLysValGlyLeuLeGly 137
Db 120 GTGTGCTCTGTGTTGGCTGCGAGATCATCGAAGTGGGATTAATGTTGGACTTATGGC 179
QY 138 ProAspLeuProPheThrAsnAsnTyrglyValTrpGluAspGluPheLysAspLeuGly 157
Db 180 CCGGATCTCCCTTTCACAAACATATATGGTGTGGGAAGATGAATTTAGAGATCTTGA 239
QY 158 LeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrlLeuAspAsnAla 177
Db 240 CTTGAAGGGGTATCGAACATGCTCGAGACACACAGTGTATATATTCATGAATGNA 299
QY 178 ProValLeulleGlyArgAlaTyrglyArgValSerArgHisLeuLeuHisGluLeu 197
Db 300 CCCATCTTGATTTGGTCTGTATGACGAGGTAGTCGACATTTGCTCATGAAGAATTA 359
QY 198 LeuLysArgCysValGluSerGlyValSerTyrlLeuAspSerLysValGluArgIleThr 217
Db 360 TTAAGAAGGTGTCTGAGTCAGGTGTTTCATATCTTAGCTCAAAAGTGGAAAGCATACG 419
QY 218 GluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArg 237
Db 420 GAATCTACCAAGTGTCTATCGCTTTGACGCTTGTGAACATGATATGATTTGCCCTGCAGG 479
QY 238 LeuAlaThrValAlaSerGlyAlaLaserGlyLysLeuLeuGluTyrglyValGlyCly 257
Db 480 CTTGCTACTGTGTGCTCTCGAGCAGCATCAGGAAGCTATTGGAATATGAGTGGGGGT 539
QY 258 ProArgValCysValGlnThrAlaTyrglyValGluValGluValGluAsnAsnProTyrl 277
Db 540 CCCAAGTTCTGTCACAAACAGCTTATGGTGTGAGGTGTAGGTGGAATAATATCCATAT 599

[illegible]

```
QY 356 -----GluTrpSer 358
Db 2529 ACATAGAACTCTTATTGCTTTTAAATTCGTTTAAATATAAACTGTAGGAGTGGTCC 2588
QY 359 TyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAla 378
Db 2589 TATATCCAGTGTGGTGGTCTCTGCAACACCGAACAAGAAATCTCGCTTGGTGTCT 2648
QY 379 AlaAlaSerMetValHisProAla----- 386
Db 2649 GCCGCTAGCATGTATACATCCGCAACAGAGTAGTAAGACGACAGACGCTTGACCCGATGTA 2708
QY 387 -----Thr-G1 388
Db 2709 CTTTATTACATCTTGAATCTGTACACTCTGACTTGTAGCTTTCCTCCCAATTAACACAGG 2768
QY 388 yTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysI1 408
Db 2769 CTATTCAGTGTGAGATCTTCTGAGCTCAAAATATGATCAGTCATCGCAGAGAT 2828
QY 408 eLeuLysGlnAspAsnSerAlaTyValValSerGlyGlnSerSerAlaValAsnIleSe 428
Db 2829 ACTAGAGAGAGAGACTACCAACAGATCAACAGT-----ANATATTC 2870
QY 428 rMetGln----- 430
Db 2871 AAGACAAGGTGAGCTCTATATAAACCAAGTTAACTATGCTGCTTTTCACACAATGA 2930
QY 431 -----AlaTrpSerSerLeuTrpProLy 438
Db 2931 GAACTTTTAAACTGTGATTTTCTATTTTGGCAGCTTGGGATACCTTATGGCCACC 2990
QY 438 sGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuValGlnLeuAs 458
Db 2991 AGAAGGAGAAAGACAGAGAGATCTTCTCTTGGCTTGCACATCATAGTTCAATTCGA 3050
QY 458 pileGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrp----- 475
Db 3051 TACCAAGGAGATTAGAAGCTTCTCCGCTACTTCTCCGCTTCCAAATGGAACGACAG 3110
QY 475 ----- 475
Db 3111 ACATGATTCGTGATTCATCAACAACTGTACTGTAAACAGAAATGTATCTTCTTTTCA 3170
QY 476 -----MetTrpTrpGlyPheLeuGlySerLeuSerSerPheAspLeuValIle 492
Db 3171 CTTACTCAGATGTGGCAAGGTTCTTAGGATCAACATTAACATCAGGAGATCTCTGTCT 3230
QY 492 uPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgH1 512
Db 3231 CTTTGTCTTATACATGCTGCTATTCACCAACAAATTTGAGAAAAGGTCTCATCAATCA 3290
QY 512 sLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
Db 3291 TCTCATCTCTGTATCCAAACCGGAGCAACCATGATGATAAAACCTATCTCAAA 3339
RESULT 15
AB024035
LOCUS
DEFINITION
Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MHM17.
ACCESSION
AB024035 BA000015
VERSION
AB024035.1 GI:4519194
KEYWORDS
Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
clone:MHM17.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (sites)
REFERENCE
Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E., Kotani,H.
and Tabata,S.
```

TITLE

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source

exon

CDS

CDS

Structural analysis of Arabidopsis thaliana chromosome 5. X.
Sequence features of the regions of 3,076,755 bp covered by sixty
P1 and TAC clones
DNA Res 7 (1), 31-63 (2000)
20181125
2 (bases 1 to 78423)
Nakamura,Y.
Direct Submission
Submitted (24-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0612, Japan (E-mail:ynakamura@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd-graph.cgi?c=MHM17
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (S.M. Hebsgaard, et al., http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://genome.wustl.edu/eddy/NetGene2-SE/),
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MP10 and the 3' clone is MUL3.
Location/Qualifiers
1..78423
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MHM17"
/clone_lib="Mitsui P1"
complement(1..30)
/note="CDS is reported in Acc# AB020747
contains similarity to protein kinase
gene_id:MP10.5"
/number=1
/evidence=not_experimental
complement(join(4221..4349,4616..4761,4851..4944,
5023..5307,5397..5821,6091..6208,6305..6432,6529..6638,
6709..6982,7070..7141))
/note="dbj|BAA91947.1
gene_id:MHM17.1
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BA097020.1"
/db_xref="GI:8777430"
/translation="MAPIRLCGDPLGRINOLFRRVQSVSKSAGPFDALICVGGFFPD
SPEILDFDYVGRQVPIPTFTGDDYGVAPKLTSTSKAENQFQKMGLEVCN
LFLRGSGKFLSLGLSVAYLSGRSSSEQCKYQDDVDALRALEPFGGCTINWFA
GVTNRAAVSDIPVGLSDSCSDSTVSELVMEVKRYHAGSNGVYFAEPTVNAETH
VTRFTGLAQVGNKKQKFLHALSTPTPTSTSPALSAKPKPTLTPYLNQDGAESKK
RPNVSDSDQWRIDVPRKQSGSGQKLCFKVCSGCPGDCHEFHQHNAREOCR
RGVLDLLIKGCKEPCSYKHEFDQESSIQRPSENANRSEKCEFCILSSPSVESH
LTVSGESFYCALPKGLVEDHILIPTEHLPNTLVSPVESLSYQNGLRNCYKD
QNDVAFVFLSKRVSHANLQVVPVPSRRALLPNIFSLAAKLGFLVTKPKNDSTYD
GRKYLQKEENALGLFVYELPDGVLVSHTEENEFVPAQGREVLGLLIPKIPDRWR
NCKLSQEEAKLAEDFKQFQFPCQ"
join(7467..7643,7741..8080,8177..8230,8376..8479)
/note="gene_id:MHM17.2
pir||C/1422
similar to unknown protein"
/codon_start=1
/evidence=not_experimental

```

/protein_id="BAA97021.1"
/db_xref="GI:8777431"
/translation="MADSVLKEVNGCRPEKIPKLDKACGSKSWKHLKLGNEVDEDEY
KROYCLFHYEFKSGEGFTWDEKRYDMFHLRPLENSPPTSIDRTNADVIQVLTFAIE
KNEAGSKLVFVHEVYSAFKANGLLCWLTFWATDMASAPTSQIQVVELWRGQDF
EIPFPRKDEMDDDVEKPPSPMFDYDDKPPVVFVRAAPEDGVFVDFRGALED
LRSGI"
CDS
  join(9471..9653,9736..10236)
  /note="gene_id:MHM17.3
  pIr||C71422
  similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAA97022.1"
/db_xref="GI:8777432"
/translation="MADSLIMEVNYRRPEKIPKLDCEGSESSSKORVVKWGTAEED
EYLRLYLFHYOFQKQFSGSIKWEQDYNFRSRMDKSPNEISERKSNVLIEMILT
AIDKNEAGTKLVFVHEVYANQLRGLTCWLTFWATDMASPPSKYIQAHWRRG
QNFHTFIFRKLPTDEIEAVEVOPSPMLYELDKPPIVFSRGPDEALPGVPVFNR
TGAGLDPDW"
CDS
  complement(join(12420..13157,13254..13394,13492..13542,
  13628..13759,13862..13927,14005..14747,14880..15036))
  /note="gene_id:MHM17.4
  unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAA97023.1"
/db_xref="GI:8777433"
/translation="MESSVAPPHRTHLPNRYRSLYHFCSSDFQREPOISLATPAV
LQELVLRTEIAESPGETCEPPENLITSEKLGNGVSGDSGGERVETISQKSLMIGDI
CNATIQDASVVRSHRDFDFELMNETQDSVPESCINFLFALDNDYDQVNLVK
PQIDQVQVSEKKAEEVPSKVESNEVISSGVLEACNGTVOREMELEKPKVDNSPV
LVDVSRIVGGDDVEGEISQDNDMLVEDDVERHEEYQVSGDTGNSHLTSHKS
FGVEVMNDQAKKIDQTSNEAKMDPGTSIKRSAPSDAKAKRAKAKRAQRI
ALGVKKLKPVPKPKIKYCEHYLKGCHGDKCFESHDTIPETKSCPCCFATQS
CMKGDCCPDHLSKXPCANFYIKGFCRGDSCIFSHKGPQASDTPSANVTVSSTK
ITAAESPDPKQSVDAIAKLPQIARVSSVAFKPFSSHSNQRNSDASSKINE
HYTPQVPPRLPSPAPKMSFLSKDQEDTVKASSAKSPNTDSDQTLKQSQOG
SFLPQPPKGISFLSPASEEQLTNREPOKPPASKNLKTTPSSHISLISAMKLAEE
FSAKVERGNNDPTEAVNKNVTVDVATRNNGNISSKILEFLSSFSHGKN"
  join(15350..15482,15570..15615,16443..16671)
  /note="gene_id:MHM17.5"
/codon_start=1
/evidence-not_experimental
/product="30S ribosomal protein S16"
/protein_id="BAA97024.1"
/db_xref="GI:8777434"
/translation="MVVRIILSRFGCKNRPFFRYMAADSRSPRDKHLEVLGFENLPL
GDQGRMGLKDRIKYILWSVGAQSPDPVORLLFRSGLLPPPPMVAWGRKGADTRP
VDPMTGRYVDAENKTVNANDNQPKEDTEAKS"
  join(17089..17134,17217..17260,17389..17484,17694..17786,
  17919..18006,18095..18156,18245..18429,18562..18739,
  18827..18870,19017..19218,19684..19770)
  /note="gene_id:MHM17.6"
/codon_start=1
/evidence-not_experimental
/product="nucleosome assembly protein"
/protein_id="BAA97025.1"
/db_xref="GI:8777435"
/translation="MSNDKDSFNVDLTSALDKDEDRAGLVNLKLNKQNLQAGHSDVL
ENITKIRRVLRVLEIQGHDEIEFKFREAALEAKYOKLYQPLYNKRYEIVNAGT
EVEGAPAKMDQGDKEATKRGVPSFWLTKMNDVISEITERDEGALYLYLKDVKW
CKIEEPKGLFEFFDQNPYFKNTLLTKAYHIDEDEPLEKRAIGTEIDWYPKCLTQ
KILKKPKGAKNAKPIITKTECESFFNFNPQVDDDDIDEERAEELQNLMEODY
DIGSTIREKLIIPHAWSWFTGEALGEEFEEDDNDDEDDEDEDEDEDEDEDEDE
DREEVSKTKKFEIILQKGRPVQVDDQGERPCKQQ"
  complement(join(23174..23239,23666..24049,24174..24863,
  25283..25388,25491..25645))
  /note="gb|AAC78547.1
  gene_id:MHM17.7
  similar to unknown protein"
/codon_start=1
/evidence-not_experimental

```

```

/protein_id="BAA97026.1"
/db_xref="GI:8777436"
/translation="MMHLILCSLYLISMDGVYNEASEPSSSSSSGSLARSILFHEYRO
SVPLQNGHVSFMAFMNKLPEYRPOESQRLAFNDTRFYQMKIEASLEWPEDEP
NRKSPANSOYLRPHYPSSSSSSSLSPNNISEYSSLLPLPKPSTTTTAVNVPVLP
LAPINMHQOEPLFRNRQREEMTOAILAVLTGSPSPSTSPQKRGATAFKR
YYSMSDRGRAPLPSVRKQSMTRAMSFNRLNINQREFTRENATTHGEGSGGG
GRYTSGPSATQLOHMIKERKREKLSLPPCTKDKASVLSIAREQSSSL
QGEISKLERNREVEAKLAREIENDLRPERFNVRHRIPESTRERLTLRVLR
GDITRVDLMLRLEFLKQIINNVSISTEARTLARAEGDTSIVLISLRLEGEWDE
SAFEAVRRVVADLAH"
CDS
  complement(join(29415..29718,29813..30081,30183..30443,
  30529..30656,32107..32716))
  /note="gene_id:MHM17.8"
/codon_start=1
/evidence-not_experimental
/product="cytokinin oxidase"
/protein_id="BAA97027.1"
/db_xref="GI:8777437"
/translation="MASYNLRSQVRLIAITIVIIITLSTPTTNTSPQWNILSHNEF
AGHSHRGOASAKDGVVNMRSVMNRDGIKVSRTCLYVDVAAMWIEVLAKTLLEGL
KHTSSSSSVESAATDFGHVTKIPPSAVLIPSSVEDITDLIKLSDQSPPLAARG
TPVSWTDYLYLTVGGLSNGSIGQTFERYQPIYNLEMDVITGKEGATATCKDNMSD
LFFAVLGGELGQFGIITRARIKLEAPRAKRWLFYIDFSETRDQERIVSKTDGVDF
LEGSIMVDHGPPDNMRSTYPPSDHLRIASVKKRHRVYCLYEVKRYDETQYTVNEE
MEELSDSLNHRVGFMYEKDVTYMDFLNRVTCGLNLSKQMDVPHVNLNLFVPKTOI
SKPDDGVFKGIILRNNTISGPVLVPMNRKNWDRMSAAIPEEDVYVAVGLRSAGFD
NWEAFDOENMEILAFCEADANNVGIQYLPYHSSQEGVWRHFGPRWNIVFVKYKYDPKM
ILSPQNIQFKINSS"
  complement(39239..39315)
  /product="tRNA-Val"
  /note="codon recognized: GUG; gene_id:MHM17.9"
  /evidence-not_experimental
  complement(join(41045..41748,41932..42184))
  /note="gene_id:MHM17.10
  pIr||T04268
  similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAA97028.2"
Alignment Scores:
Pred. No.: 1,54e-128 Length: 78423
Score: 1633.50 Matches: 376
Percent Similarity: 53.36% Conservativity: 61
Best Local Similarity: 45.91% Mismatches: 80
Query Match: 59.55% Indels: 303
DB: 8 Gaps: 15
US-09-701-395A-23 (1-529) x AB024035 (1-78423)
Qy 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTrrp 16
Db 61869 ATGGAGTGTGTGGGGCTAGGAATTTCCGACCAATGGCGGTTCACATTCCTCCGTCATGG 61928
Qy 17 ThrPheGlyThrArgAsnLeuSerSerSerLeuAlaTyrAsnIleHisArgTyrGly 36
Db 61929 AGTTGTCGAAGGAATTTCCAGTGTGTTAAAGATACATAGCAATATTCGTTTCGCT 61988
Qy 37 SerSerCysArgValAspPheClnValArgAlaAspGlyGly-----SerGlySerArg 54
Db 61989 --TTGTGT-----AGTGTACAGCTAGCGCGGGAAGTTCCGTAGTAGAG 62033
Qy 55 SerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
Db 62034 AGTTGTAGCGGTGACGAGAGAGATTCCCTGACGAAGAAGATTGTGTAACACTGTGTGT 62093
Qy 75 SerGluLeuLeuPheValClnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
Db 62094 TCTGAGATTCTATTGTTCAATGCGACAGACAAGATATGATGATGACAGCTTAAGCTT 62153
Qy 95 AlaAspLys----- 97

```

[illegible]

Search completed: May 21, 2003, 23:41:27
Job time : 4376 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 22:11:13 ; Search time 189 Seconds
(without alignments)
277.579 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRNLSSCPVWTRGT.....VRHLSDPSGAVMVRAYLER 529

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172655 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2743	100.0	529	10	US-09-323-998D-23
2	2743	100.0	529	10	US-09-323-998D-50
3	2737	99.8	529	10	US-09-323-998D-47
4	2728	99.5	529	10	US-09-323-998D-51
5	1959.5	71.4	524	10	US-09-323-998D-2
6	1959.5	71.4	524	10	US-09-323-998D-21
7	1959.5	71.4	524	10	US-09-323-998D-48
8	1934	70.5	516	10	US-09-323-998D-52
9	1934	70.5	533	10	US-09-323-998D-54
10	1909.5	69.6	526	10	US-09-323-998D-53
11	1882	68.6	533	10	US-09-323-998D-26
12	1518.5	55.4	378	10	US-09-323-998D-25
13	1518.5	55.4	378	10	US-09-323-998D-47
14	1513	55.2	374	10	US-09-323-998D-27
15	775.5	28.3	502	10	US-09-323-998D-56
16	761	27.7	498	10	US-09-323-998D-57
17	755	27.5	511	10	US-09-323-998D-60
18	749.5	27.3	500	10	US-09-323-998D-58
19	748	27.3	501	10	US-09-323-998D-55

20	737	26.9	500	10	US-09-323-998D-59	Sequence 59, Appl
21	701.5	25.6	456	10	US-09-323-998D-20	Sequence 20, Appl
22	701.5	25.6	503	10	US-09-323-998D-61	Sequence 61, Appl
23	129.5	4.7	382	9	US-09-941-947A-30	Sequence 30, Appl
24	115.5	4.2	424	9	US-09-738-626-4029	Sequence 4029, Ap
25	113	4.1	382	9	US-09-920-923-5	Sequence 5, Appl1
26	113	4.1	382	10	US-09-547-267-7	Sequence 7, Appl1
27	103	3.8	433	10	US-09-815-242-5893	Sequence 5893, Ap
28	103	3.8	574	9	US-10-025-380-1107	Sequence 1107, Ap
29	103	3.8	574	10	US-09-922-217-1107	Sequence 1107, Ap
30	101	3.7	435	10	US-09-815-242-10696	Sequence 10696, A
31	99.5	3.6	567	10	US-09-815-242-13563	Sequence 13563, A
32	98	3.6	392	10	US-09-815-242-11795	Sequence 11795, A
33	98	3.6	435	10	US-09-815-242-12859	Sequence 12859, A
34	96.5	3.5	1146	10	US-09-824-734-2	Sequence 2, Appl1
35	95	3.5	419	10	US-09-815-242-13798	Sequence 13798, A
36	93.5	3.4	722	9	US-09-906-419-19	Sequence 19, Appl
37	93	3.4	469	9	US-09-738-626-4264	Sequence 4264, Ap
38	92	3.4	392	10	US-09-924-256A-24	Sequence 24, Appl
39	92	3.4	548	9	US-09-738-626-6486	Sequence 6486, Ap
40	92	3.4	751	10	US-09-815-242-5832	Sequence 5832, Ap
41	92	3.4	825	10	US-09-815-242-12963	Sequence 12963, A
42	91.5	3.3	799	10	US-09-952-677-6	Sequence 6, Appl1
43	89.5	3.3	1140	9	US-10-108-605-293	Sequence 293, App
44	89	3.2	400	10	US-09-815-242-10380	Sequence 10380, A
45	89	3.2	451	10	US-09-815-242-13127	Sequence 13127, A

ALIGNMENTS

RESULT 1

US-09-323-998D-23

Sequence 23, Application US/0932323998D

Patent No. US20020102631A1

GENERAL INFORMATION:

APPLICANT: CUNNINGHAM JR., FRANCIS X.

APPLICANT: SUN, ZAIREN

TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND

TITLE OF INVENTION: METHODS OF USE THEREOF

FILE REFERENCE: 108172-09019

CURRENT APPLICATION NUMBER: US/09/323,998D

CURRENT FILING DATE: 1999-06-02

PRIOR APPLICATION NUMBER: 09/088,724

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 09/088,725

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 08/937,155

PRIOR FILING DATE: 1997-09-25

PRIOR APPLICATION NUMBER: 08/624,125

PRIOR FILING DATE: 1996-03-29

NUMBER OF SEQ ID NOS: 61

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 23

LENGTH: 529

TYPE: PRT

ORGANISM: Adonis palaeatina

US-09-323-998D-23

Query Match 100.0%; Score 2743; DB 10; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.1e-247;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVRNLSSCPVWTRGTFRNLSSKLAYNHRYGSSCRVDFQVRADGGSSRSVAYK 60

Db 1 MELLGVRNLSSCPVWTRGTFRNLSSKLAYNHRYGSSCRVDFQVRADGGSSRSVAYK 60

QY 61 EGVFDEDFTKAGSELLFQVMOOTKSMKQAKLADLPPIPGESVMDLVVIGCGPAGL 120

Db 61 EGVFDEDFTKAGSELLFQVMOOTKSMKQAKLADLPPIPGESVMDLVVIGCGPAGL 120

QY 121 SLAAEAAKLGKVLGIPDLPTFTNNYQWDEDFKDLGLERCIEHAWKDTTVYLDNDAPVL 180

Db 121 SLAAEAAKLGKVLGIPDLPTFTNNYQWDEDFKDLGLERCIEHAWKDTTVYLDNDAPVL 180

Db 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLAT 240
QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLVFMVDYRDYMOOKLOCSEE 300
Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLVFMVDYRDYMOOKLOCSEE 300
QY 301 EYPTFLYVMPMSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
Db 301 EYPTFLYVMPMSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
QY 361 PVGGLPNTQKNLAFGAASVHPATGYSVVRSLSSEAPKYASVIAKILKQNSAYVYSG 420
Db 361 PVGGLPNTQKNLAFGAASVHPATGYSVVRSLSSEAPKYASVIAKILKQNSAYVYSG 420
QY 421 QSSAVNISQAWSSLPWPKERKQRAFFLGLLEIYVOLDIEATRTFFRFLPTMMWNGF 480
Db 421 QSSAVNISQAWSSLPWPKERKQRAFFLGLLEIYVOLDIEATRTFFRFLPTMMWNGF 480
QY 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSCGAVMVRAYLER 529
Db 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSCGAVMVRAYLER 529

RESULT 2

US-09-323-998D-50
; Sequence 50, Application us/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:

; APPLICANT: CUNNINGHAM JR., FRANCIS X.

; APPLICANT: SUN, ZAIREN

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND

; FILE REFERENCE: 108172-09019

; CURRENT APPLICATION NUMBER: US/09/323,998D

; CURRENT FILING DATE: 1999-06-02

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1997-09-25

; PRIOR FILING DATE: 1996-03-29

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 50

; LENGTH: 529

; TYPE: PRT

; ORGANISM: Adonis palaestina

US-09-323-998D-50

Query Match 100.0%; Score 2743; DB 10; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.1e-247;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVRNLISCCPVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGRSVAYK 60
Db 1 MELLGVRNLISCCPVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGRSVAYK 60
QY 61 EGFVDEEDFKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
QY 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
Db 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLAT 240

Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLAT 240
QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLVFMVDYRDYMOOKLOCSEE 300
Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLVFMVDYRDYMOOKLOCSEE 300
QY 301 EYPTFLYVMPMSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
Db 301 EYPTFLYVMPMSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
QY 361 PVGGLPNTQKNLAFGAASVHPATGYSVVRSLSSEAPKYASVIAKILKQNSAYVYSG 420
Db 361 PVGGLPNTQKNLAFGAASVHPATGYSVVRSLSSEAPKYASVIAKILKQNSAYVYSG 420
QY 421 QSSAVNISQAWSSLPWPKERKQRAFFLGLLEIYVOLDIEATRTFFRFLPTMMWNGF 480
Db 421 QSSAVNISQAWSSLPWPKERKQRAFFLGLLEIYVOLDIEATRTFFRFLPTMMWNGF 480
QY 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSCGAVMVRAYLER 529
Db 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSCGAVMVRAYLER 529

RESULT 3

US-09-323-998D-47
; Sequence 47, Application us/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:

; APPLICANT: CUNNINGHAM JR., FRANCIS X.

; APPLICANT: SUN, ZAIREN

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND

; FILE REFERENCE: 108172-09019

; CURRENT APPLICATION NUMBER: US/09/323,998D

; CURRENT FILING DATE: 1999-06-02

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1997-09-25

; PRIOR FILING DATE: 1996-03-29

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 47

; LENGTH: 529

; TYPE: PRT

; ORGANISM: Adonis palaestina

US-09-323-998D-47

Query Match 99.8%; Score 2737; DB 10; Length 529;
Best Local Similarity 99.6%; Pred. No. 4.1e-247;
Matches 527; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVRNLISCCPVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGRSVAYK 60
Db 1 MELLGVRNLISCCPVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGRSVAYK 60
QY 61 EGFVDEEDFKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
QY 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
Db 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLAT 240
QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLVFMVDYRDYMOOKLOCSEE 300
Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLVFMVDYRDYMOOKLOCSEE 300

Db 241 VASGAASGLLEYEVGGPRVCQTAYGVEVEENNPDNLMVMDYRDYMOOKLOCSEE 300
QY 301 EYPTFLYVMPMSPTLRFEEETCLASDAMPFDLLKRLKSLKTLGIVTKYVEEWSYI 360
Db 301 EYPTFLYVMPMSPTLRFEEETCLASDAMPFDLLKRLKSLKTLGIVTKYVEEWSYI 360
QY 361 PVGGSPLNTEQKNLAFAGAAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVVG 420
Db 361 PVGGSPLNTEQKNLAFAGAAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVVG 420
QY 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPMTWMMWGF 480
Db 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPMTWMMWGF 480
QY 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVYAYLER 529
Db 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVYAYLER 529

RESULT 4

US-09-323-998D-51
; Sequence 51, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Adonis palaestina
US-09-323-998D-51

Query Match 99.5%; Score 2728; DB 10; Length 529;
Best Local Similarity 99.1%; Pred. No. 2.9e-246;
Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELLGVRLNLISSCPVMTFGRNLSSKLAYNTHRYGSSCRVDFQVRADGSGSRSSVAYK 60
Db 1 MELLGVRLNLISSCPVMTFGRNLSSKLAYNTHRYGSSCRVDFQVRADGSGSRSSVAYK 60
QY 61 EGFVDEEDFIKAGGSELLFVQMQTKSMEKQAKLADKLPPFPGESVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFIKAGGSELLFVQMQTKSMEKQAKLADKLPPFPGESVMDLVVIGCGPAGL 120
QY 121 SLAAEAAKGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIYLDNDAPVL 180
Db 121 SLAAEAAKGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIPICRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIPICRLAT 240
QY 241 VASGAASGLLEYEVGGPRVCQTAYGVEVEENNPDNLMVMDYRDYMOOKLOCSEE 300
Db 241 VASGAASGLLEYEVGGPRVCQTAYGVEVEENNPDNLMVMDYRDYMOOKLOCSEE 300
QY 301 EYPTFLYVMPMSPTLRFEEETCLASDAMPFDLLKRLKSLKTLGIVTKYVEEWSYI 360

Db 301 EYPTFLYVMPMSPTLRFEEETCLASDAMPFDLLKRLKSLKTLGIVTKYVEEWSYI 360
QY 361 PVGGSPLNTEQKNLAFAGAAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVVG 420
Db 361 PVGGSPLNTEQKNLAFAGAAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVVG 420
QY 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPMTWMMWGF 480
Db 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPMTWMMWGF 480
QY 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVYAYLER 529
Db 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVYAYLER 529

RESULT 5

US-09-323-998D-2
; Sequence 2, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-323-998D-2

Query Match 71.4%; Score 1959.5; DB 10; Length 524;
Best Local Similarity 70.4%; Pred. No. 1.9e-174;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
QY 1 MELLGVRLNLISSCPVMTFGRNLSSKLAYNTHRYGSSCRVDFQVRADGSGSGSR 54
Db 1 MECVGARNEAAMAVSTFPSSCRKRKPPVVKRYSYRNIRFG-LC---SVRASGGSGSGSE 55
QY 55 SVYAYKEGFVDEEDFIKAGGSELLFVQMQTKSMEKQAKLADKLPPFPGESVMDLVVIG 114
Db 56 SCYAVRDEFADEEDFIKAGGSELLFVQMQTKSMEKQAKLADKLPPFPGESVMDLVVIG 115
QY 115 CGPAGLSLAAEAAKGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIYLD 174
Db 116 CGPAGLSLAAEAAKGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIYLD 175
QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIPIC 234
Db 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIPIC 235
QY 235 PCRLATVASGAASGLLEYEVGGPRVCQTAYGVEVEENNPDNLMVMDYRDYMOOK 294
Db 236 PCRLATVASGAASGLLEYEVGGPRVCQTAYGVEVEENNPDNLMVMDYRDYMOOK 295
QY 295 LOCSESEYPTFLYVMPMSPTLRFEEETCLASDAMPFDLLKRLKSLKTLGIVTKYVE 354
Db 296 VRSLEAEYPTFLYVMPMSPTLRFEEETCLASDAMPFDLLKRLKSLKTLGIVTKYVE 355
QY 355 EWSYIPVGGSLPTEQKNLAFAGAAASVHPATGYSVVRSLSEAPKYASIAKILKQDNS 414

Db 356 EWSYIPVGGSLPNTQKKNLAFGAASVMHPATGYSVRSLSSEAPKYASVIAEILREETT 415
QY 415 AYYVSGSSAVNISQAWSSLPKRRQRAFFLGLGLIVOLDIEATFTFFRFLPT 474
Db 416 KQINS-----NISQAWDTLWPPERKQRAFFLGLGLIVOLDIEATFTFFRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFMTMFLVAPNSMRSLVRHLLSDPSGAVMVRAYLE 528
Db 470 WNWGFLGSLTSGDLVLFALYMFVISPNNLRKGLNHLISDPTGATMIKTYLK 523

RESULT 6
US-09-323-998D-21
; Sequence 21, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1997-09-25
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-323-998D-21

Query Match 71.4%; Score 1959.5; DB 10; Length 524;
Best Local Similarity 70.4%; Pred. No. 1.9e-174;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
QY 1 MELLGVRLN-----ISSCPVMTFTGRNLSSSKLAAYNIHRYGSSCRVDFQVRADGG--SGSR 54
Db 1 MECVGARNFAAMAVSTPFSCKRKFPPVKRYSTNRIRFG--LC-----SVRASGGSSGSE 55
QY 55 SSVAYKEGFVDEEDFIKAGGSELLFVQMOKTSMKQAKLADKLPPIPFGESVMDLVVIG 114
Db 56 SCVAVREDFADEEDFVKAGGSEILFVQMOKMDQSKLVKLPPIISIGDGLDHHVIG 115
QY 115 CGPAGLSLAAEAAGLGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 174
Db 116 CGPAGLALAAESAKLGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 175
QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGLRLVACDDNNVI 235
QY 235 PCRLATVASGAAGSKLLEYEVGGPRVCVQATAYGVEVEVNNPYDPNLMFMDYRDYMOOK 294
Db 236 PCRLATVASGAAGSKLLEYEVGGPRVCVQATAYGVEVEVNNPYDPNLMFMDYRDYMOOK 295
QY 295 LOCSEEEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVQVTKYE 354
Db 296 VRSLEAEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVQVTKYE 355
QY 355 EWSYIPVGGSLPNTQKKNLAFGAASVMHPATGYSVRSLSSEAPKYASVIAEILREETT 414
Db 356 EWSYIPVGGSLPNTQKKNLAFGAASVMHPATGYSVRSLSSEAPKYASVIAEILREETT 415
QY 415 AYYVSGSSAVNISQAWSSLPKRRQRAFFLGLGLIVOLDIEATFTFFRFLPT 474
Db 416 KQINS-----NISQAWDTLWPPERKQRAFFLGLGLIVOLDIEATFTFFRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFMTMFLVAPNSMRSLVRHLLSDPSGAVMVRAYLE 528

Db 416 KQINS-----NISQAWDTLWPPERKQRAFFLGLGLIVOLDIEATFTFFRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFMTMFLVAPNSMRSLVRHLLSDPSGAVMVRAYLE 528
Db 470 WNWGFLGSLTSGDLVLFALYMFVISPNNLRKGLNHLISDPTGATMIKTYLK 523

RESULT 7
US-09-323-998D-49
; Sequence 49, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1997-09-25
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-323-998D-49

Query Match 71.4%; Score 1959.5; DB 10; Length 524;
Best Local Similarity 70.4%; Pred. No. 1.9e-174;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
QY 1 MELLGVRLN-----ISSCPVMTFTGRNLSSSKLAAYNIHRYGSSCRVDFQVRADGG--SGSR 54
Db 1 MECVGARNFAAMAVSTPFSCKRKFPPVKRYSTNRIRFG--LC-----SVRASGGSSGSE 55
QY 55 SSVAYKEGFVDEEDFIKAGGSELLFVQMOKTSMKQAKLADKLPPIPFGESVMDLVVIG 114
Db 56 SCVAVREDFADEEDFVKAGGSEILFVQMOKMDQSKLVKLPPIISIGDGLDHHVIG 115
QY 115 CGPAGLSLAAEAAGLGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 174
Db 116 CGPAGLALAAESAKLGLKVLGIGLDPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 175
QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGLRLVACDDNNVI 235
QY 235 PCRLATVASGAAGSKLLEYEVGGPRVCVQATAYGVEVEVNNPYDPNLMFMDYRDYMOOK 294
Db 236 PCRLATVASGAAGSKLLEYEVGGPRVCVQATAYGVEVEVNNPYDPNLMFMDYRDYMOOK 295
QY 295 LOCSEEEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVQVTKYE 354
Db 296 VRSLEAEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVQVTKYE 355
QY 355 EWSYIPVGGSLPNTQKKNLAFGAASVMHPATGYSVRSLSSEAPKYASVIAEILREETT 414
Db 356 EWSYIPVGGSLPNTQKKNLAFGAASVMHPATGYSVRSLSSEAPKYASVIAEILREETT 415
QY 415 AYYVSGSSAVNISQAWSSLPKRRQRAFFLGLGLIVOLDIEATFTFFRFLPT 474
Db 416 KQINS-----NISQAWDTLWPPERKQRAFFLGLGLIVOLDIEATFTFFRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFMTMFLVAPNSMRSLVRHLLSDPSGAVMVRAYLE 528

Db 470 WMWQFGLSTLTSGDLVLFALYMFVISPNNLRKGLINHLSIPTGATMIKTYLK 523

RESULT 8

```

US-09-323-998D-54
; Sequence 54, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatenIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Tagetes erecta
US-09-323-998D-54

```

RESULT 9

```

US-09-323-998D-52
; Sequence 52, Application US/093232998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Lactuca sp.
US-09-323-998D-52

```

RESULT 10

US-09-323-998D-53
; Sequence 53, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 53
; TYPE: PRT
; LENGTH: 526
; ORGANISM: Lycopodium esculentum
US-09-323-998D-53

Query Match 69.6%; Score 1909.5; DB 10; Length 526;
Best Local Similarity 68.5%; Pred. No. 8.9e-170;
Matches 368; Conservative 68; Mismatches 78; Indels 23; Gaps 6;
QY 1 MELLGVRNLSSCPV-----WTFGRNLSSK---LAYNIHRYGSSCRVDFOVRADGG 50
DB 1 MECVGQVNGVAMVLTPLRLNRWGG--ELCQKSIPLAY--EQYESKC-----NSS 48
QY 51 SGRSSVAYKEGFDEDFIKAGSELFFVQMOTKMEKAKLADKLPLPPFGESVMDL 110
DB 49 SGDSVCVVDKEDFADEEDYIKAGSQLVFVQMOKKMDQOQKSLDELQISAGQTVL 108
QY 111 VVIGCGPAGLSAAEAAKGLKVLGPDLPFTNNYGVWEDEFKDLGLERCIHAWKDTI 170
DB 109 VVIGCGPAGLALAAESAKGLNVLGVPDLPFTNNYGVWEDEFKDLGLQACIEHWRTI 168
QY 171 VYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVESGVYLSKVERITEAGDGLSVVCEN 230
DB 169 VYLDNDPELIGRAYGRVSRHLLHEELLKRCVEAGVLYLSKVDRIVEATNGQSLVCEG 228
QY 231 EIFTPCRLATVASGAASKLEVEVGGPRVCQTAYGVEVEVNNPDNLMVFMVMDYRDY 290
DB 229 DVVTPCRFVTVASGAASKFLQELGSPRVSVQVAYGVEVEVNNPDNLMVFMVMDYRDY 288
QY 291 MQOKLQCSSEYPTFLYVMPSPTRLPFEETCLASKDAMPFDLLKRLMSRLKTLGIQVT 350
DB 289 LRHDAQSLEAKYPTFLYAMPSPTRVFEETCLASKDAMPFDLLKRLMLRLNLTGLVRIK 348
QY 351 KYVEEWSYIPVGSGLPNTQKNLAFGAAASMVHPATGYSVWRSLSAPYASVIKILK 410
DB 349 EYVEEWSYIPVGSGLPNTQKTLAFGAAASMVHPATGYSVWRSLSAPYASVIKILK 408
QY 411 QDMSAYVVGSGSAVNISMQAWSLWPKERKQRAFFLFGLELTVQIDIEATRTFFTF 470
DB 409 QHYSKKNLT--SSSTPSISTQAWNTLWPKERKQRAFFLFGLELTVQIDIEATRTFFTF 467
QY 471 RLPTWMMWGLGSSLSFDLVLSMYNVLAPNSMRSLVRHLLSDPSGAVMVRAYL 527
DB 468 RVPKMMWQGLGSSLSADLMLFAFYMFIIAPNDMRGLRHLLSDPTGATLRTYL 524

RESULT 11

US-09-323-998D-26
; Sequence 26, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:

APPLICANT: CUNNINGHAM JR., FRANCIS X.
APPLICANT: SUN, ZAIREN
TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 108172-09019
CURRENT APPLICATION NUMBER: US/09/323,998D
CURRENT FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: 09/088,724
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 09/088,725
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 08/937,155
PRIOR FILING DATE: 1997-09-25
PRIOR APPLICATION NUMBER: 08/624,125
PRIOR FILING DATE: 1996-03-29
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 26
LENGTH: 533
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: Lactuca sp./Solanum sp.
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (491)
OTHER INFORMATION: Any amino acid
US-09-323-998D-26

Query Match 68.6%; Score 1882; DB 10; Length 533;
Best Local Similarity 67.6%; Pred. No. 3.4e-167;
Matches 370; Conservative 65; Mismatches 76; Indels 36; Gaps 8;
QY 1 MELLGVRNL-----ISSCPVMT-----FGTR--NLSSSKLAYNIHRYGSSCRVD 42
DB 1 MECVGARNMTAMVFTCPRTDCNIRHKFSLKGRFTNLAS-----SSLR-- 48
QY 43 FQVRADGSGSRSSVAYKEGF--VDEEDFIKAGSELFFVQMOTKMEKAKLADKLPP 100
DB 49 -QIKC--SAKSDRCVVDKQIGSVADVEDYKAGSELFFVQMOTKMEKAKLADKLPP 105
QY 101 IPGESVMDLVVIGCGPAGLSAAEAAKGLKVLGPDLPFTNNYGVWEDEFKDLGLER 160
DB 106 IPTGNCITLDLVVIGCGPAGLSAAEAAKGLNVLGPDLPFTNNYGVWODEFTGLGLEG 165
QY 161 CIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVESGVYLSKVERITEAG 220
DB 166 CIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVESGVYLSKVERITEAG 225
QY 221 DGHSLVVCENEIETPCRLATVASGAASKLEVEVGGPRVCQTAYGVEVEVNNPDN 280
DB 226 NGYSLICEGNIETPCRLATVASGAASKLEVEVGGPRVCQTAYGVEVEVNNPDN 285
QY 281 LMVFMVMDYRDYMQOKLQCSSEYPTFLYVMPSPTRLPFEETCLASKDAMPFDLLKRLMS 340
DB 286 LMVFMVMDYRDYRDHDAQSLEAKYPTFLYAMPSPTRVFEETCLASKDAMPFDLLKRLMS 345
QY 341 RLKTLGIQVTKVVEEWSYIPVGSGLPNTQKNLAFGAAASMVHPATGYSVWRSLSAPK 400
DB 346 RLKTLGIQVTKVVEEWSYIPVGSGLPNTQKTLAFGAAASMVHPATGYSVWRSLSAPK 405
QY 401 YASVIKILKQDMSAYVVGSGSAVNISMQAWSLWPKERKQRAFFLFGLELTVQIDIE 460
DB 406 CAFVLANILRQNSKNLT--SSSTPSISTQAWNTLWPKERKQRAFFLFGLELTVQIDIE 464
QY 461 ATFTFTFTFTPLTWMWGLGSSLSFDLVLSMYNVLAPNSMRSLVRHLLSDPSGA 520
DB 465 GRSFTRAFTRVPKMMWQGLGSSLSADLMLFAFYMFIIAPNDMRGLRHLLSDPTGA 524
QY 521 VMVRAYL 527
DB 525 TLIRTYL 531

RESULT 12

US-09-323-998D-25
; Sequence 25, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (336)
; OTHER INFORMATION: Any amino acid
US-09-323-998D-25

Query Match 55.4%; Score 1518.5; DB 10; Length 378;
Best Local Similarity 74.5%; Pred. No. 1.9e-133;
Matches 281; Conservative 48; Mismatches 47; Indels 1; Gaps 1;

QY 151 DEFKDLGIERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 210
DB 1 DEFKDLGLOACIEHVWRDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 60
QY 211 SKVERITEAGDCHSLVVCENEIFPCRLATVASGAASKLLEYEGGPRVCVQTAYGVEV 270
DB 61 SKVDRIEATNGHSLVECEGVVPCRFVTVASGAASKGLFQYELGGRVSVQTYGVEV 120
QY 271 EVENNPYPNLMVMDYRDYMOOKQCSEEEYPTFLYVMPSPTRLFFETCTCLASKDAMP 330
DB 121 EVDNPPFDSLVMFYMDYRDYVHRDAQSLEAKYPTFLYAMPSPTRVFFETCTCLASKDAMP 180
QY 331 FDLKRRKLSRLKTLGIGQVTKYVEEWSYIPVGGSLPNTPEQKNLAFGAASVMVHPATGYS 390
DB 181 FDLKRRKLSRLKTLGIRSVRIKEIYEBSYIPVGGSLPNTPEQKNLAFGAASVMVHPATGYS 240
QY 391 VVRSLEAPKVASVIAKILKQNSAVVYVSGSSAVNISQWSSLPKPKRQRAFFFLG 450
DB 241 VVRSLEAPKCAFVLANTLRQNHKNMLT-SSSTPSTISQWNTLWPKRQRAFFFLG 299
QY 451 LELIVQLDIEATRTFFRFTPLTMMWGLFGLSSSFDLVLFMSYMFVLPAPNMRMSIV 510
DB 300 LALIQLDIEGIRSFRAFFRVPKMWMQGLGSSLSXADLMFLAFYMFIIAPNDMRRGLI 359
QY 511 RHLLSDPSGAVMVRAYL 527
DB 360 RHLLSDPTGATLIRTYL 376

RESULT 13

US-09-323-998D-48
; Sequence 48, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (336)
; OTHER INFORMATION: Any amino acid
US-09-323-998D-48

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (336)
; OTHER INFORMATION: Any amino acid
US-09-323-998D-48

Query Match 55.4%; Score 1518.5; DB 10; Length 378;
Best Local Similarity 74.5%; Pred. No. 1.9e-133;
Matches 281; Conservative 48; Mismatches 47; Indels 1; Gaps 1;

QY 151 DEFKDLGIERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 210
DB 1 DEFKDLGLOACIEHVWRDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 60
QY 211 SKVERITEAGDCHSLVVCENEIFPCRLATVASGAASKLLEYEGGPRVCVQTAYGVEV 270
DB 61 SKVDRIEATNGHSLVECEGVVPCRFVTVASGAASKGLFQYELGGRVSVQTYGVEV 120
QY 271 EVENNPYPNLMVMDYRDYMOOKQCSEEEYPTFLYVMPSPTRLFFETCTCLASKDAMP 330
DB 121 EVDNPPFDSLVMFYMDYRDYVHRDAQSLEAKYPTFLYAMPSPTRVFFETCTCLASKDAMP 180
QY 331 FDLKRRKLSRLKTLGIGQVTKYVEEWSYIPVGGSLPNTPEQKNLAFGAASVMVHPATGYS 390
DB 181 FDLKRRKLSRLKTLGIRSVRIKEIYEBSYIPVGGSLPNTPEQKNLAFGAASVMVHPATGYS 240
QY 391 VVRSLEAPKVASVIAKILKQNSAVVYVSGSSAVNISQWSSLPKPKRQRAFFFLG 450
DB 241 VVRSLEAPKCAFVLANTLRQNHKNMLT-SSSTPSTISQWNTLWPKRQRAFFFLG 299
QY 451 LELIVQLDIEATRTFFRFTPLTMMWGLFGLSSSFDLVLFMSYMFVLPAPNMRMSIV 510
DB 300 LALIQLDIEGIRSFRAFFRVPKMWMQGLGSSLSXADLMFLAFYMFIIAPNDMRRGLI 359
QY 511 RHLLSDPSGAVMVRAYL 527
DB 360 RHLLSDPTGATLIRTYL 376

RESULT 14

US-09-323-998D-27
; Sequence 27, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (336)
; OTHER INFORMATION: Any amino acid
US-09-323-998D-27

; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 27
; LENGTH: 374
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
US-09-323-998D-27

Query Match 55.2%; Score 1513; DB 10; Length 374;
Best Local Similarity 74.7%; Pred. No. 6e-133;
Matches 283; Conservative 42; Mismatches 48; Indels 6; Gaps 1;
QY 150 EDEFKDLGERCTEHAWKDTIIVLDNDAPVLIGRAYGRVSRHLLHE 209
DB 1 EDEFNDLGLQKCIHVWRETIVYLDLDDKPTIGRAYGRVSRHLLHE 60
QY 210 DSKVERITEAGDGHSLVVCENEIFPCRLATVASGAAGKLLVEYGV 269
DB 61 SSKVDSITEASDGLRLVACDNNVPCRLATVASGAAGKLLVEYGV 120
QY 270 VEENNPDPNLMVMDYRDYMOQKLCSEEEYPTFLYVMPSPTRLF 329
DB 121 VEENSPYDPQVMYMDYRDYTNKVRSLAEYPTFLYAMPMTKSR 180
QY 330 PFLLKRLKMSRLKTIGIQVTKVVEEWSYIPVGGSLPNTFQKNL 389
DB 181 PFLLKRLKMSRLKTIGIRLTKVVEEWSYIPVGGSLPNTFQKNL 240
QY 390 SVYRSLEAPKYSVIAKILKODNSAYVVGSSAVNISQANSSLP 449
DB 241 SVYRSLEAPKYSVIAKILKODNSAYVVGSSAVNISQANSSLP 294
QY 450 GLELIVOLDIEATRTFFRTFFRLPTMMWGLGSSLSFSLVLF 509
DB 295 GLALIVQDFTEGIRSFRTFFRLPKMWMQGLGSLTSLTSGDL 354
QY 510 VRLLSDPSCAVMVRAYLE 528
DB 355 INHLISDPTGATMIKTYLK 373

RESULT 15
US-09-323-998D-56
; Sequence 56, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 502
; TYPE: PRP
; ORGANISM: Adonis palaestina
US-09-323-998D-56

Query Match 28.3%; Score 775.5; DB 10; Length 502;
Best Local Similarity 36.3%; Pred. No. 8e-64;
Matches 181; Conservative 99; Mismatches 166; Indels 53; Gaps 14;

QY 23 LSSSKLAYNIHRYGS-----SCRVDFOVRADGGSGSRSSVAYKEGFVDEEDFIKAGSELL 78
DB 27 VSTSKLQNOVFRIASRNHPCR-----NGT-----VKARGSALL 60
QY 79 FVQMOQTKSMKQAKLADKLPP-IPGESYMDLVVVICGPGAGLSLAAEAANKLGLVGLI- 136
DB 61 -----ELVPETKKNLEFDLPAYDPSRGIVVDLAVVGGGPGAGLAIQAQVSEAGLLVCSID 115
QY 137 -GPDLPFTNNYGVWEDEFKDLGERCTEHAWKDTIIVLDNDAPVLIGRAYGRVSRHLLHE 195
DB 116 PPSKLLWPNNYGVWDEFAMDLDDCLDTTWSGAVYTTDDNSKKYLDLDRPYGRVNRKQLKS 175
QY 196 ELKRCVESGVSYLDSKVERITEAGDGHSLVVCENEIFPCRLATVASGAAGKLLVEYEV 255
DB 176 KMLQKCVTNGVKPQAKVIKVIHE-ESKSLICNDGITTINATVVLDTATG-FSRCLVQYDK 233
QY 256 G-GPRVCVQTAYGEVEVENNPYDNLVPMFYRD-YMOOKLQCSSE--EYPTFLYVMPM 311
DB 234 PYNPG--YQVAYGIMAEVEEHPFDLKMFLMDWRDLSHLNEKLEKDKNRKIPITFLYAMPF 291
QY 312 SPTRLEFEETCLASKDAMPFDLKRKLMSRLKTIGIQVTKVVEEWSYIPVGGSLPNTFQ 371
DB 292 SSKIFLEETSLVARPGLREFDIOERVARLKHGKVKSIEDERCVPIMGGPLPVLPO 351
QY 372 KNLAFGAAASMVHPATGYSVVRSLSEAPKYSVIAKILKODNSAYVVGSSAVNISQANSS 431
DB 352 RVVGIGGTAGMVHPSTGYMVARTLAAAPVVAKSIVQVYLGSDRS---LSGN---ELSAEV 404
QY 432 WSLWPKERKQRAFFLGLLELIVOLDIEATRTFFRTFFRLPTMMWGLGSSLSFSLV 491
DB 405 WDLWPIERRRQREFFCFGMDILLKLDLQGTTRFFDAFFDLEPHYWHGFLSSRLFLPELL 464
QY 492 LFSMFMFLAPNSMRMSLV 510
DB 465 FTGLSLFSSHASNASRIEM 483

Search completed: May 21, 2003, 22:22:05
Job time : 192 secs

Result No.	Query No.	Score	Match	Length	DB	ID	Description
1	1959.5	71.4	524	1	LCYE_ARATH	Q38932	arabidopsis
2	1909.5	69.6	526	1	LCYE_LYCES	Q05837	lycopersico
3	761	27.7	498	1	LCYB_CAPAN	Q43415	capsicum an
4	749.5	27.3	500	1	LCYB_LYCES	Q43503	lycopersico
5	748	27.3	501	1	LCYB_ARATH	Q38933	arabidopsis
6	737	26.9	500	1	LCYB_TOBAC	Q43578	nicotiana t
7	708	25.8	503	1	CCS_CITSI	Q38ea0	citrus sine
8	705.5	25.7	498	1	CCS_CAPAN	Q42435	capsicum an
9	701.5	25.6	503	1	LCYB_NARPS	Q40424	narcissus p
10	602	21.9	411	1	LCYB_SYNP7	Q52776	synecococc
11	429.5	15.7	410	1	Y801_DEIRA	Q97w68	deinococcu
12	137.5	5.0	506	1	ER11_BRANA	Q65727	brassica na
13	129	4.7	386	1	CRTY_AGRAU	P54974	agrobacteri
14	125.5	4.6	382	1	CRTY_PANAN	P21687	pantoea ana
15	125	4.6	386	1	CRTY_ERWHE	Q01331	erwinia her
16	112	4.1	518	1	ER12_BRANA	Q65726	brassica na
17	111	4.0	517	1	ER12_ARATH	Q65402	arabidopsis
18	109.5	4.0	391	1	Y532_MERJA	Q57952	methanococ
19	106.5	3.9	372	1	FTZL_PYRAB	Q3v280	pyrococcus
20	105.5	3.8	414	1	YD00_SYNF3	P72835	synecocyst
21	105	3.8	539	1	ERGL_PANGI	Q48651	panax gins
22	105	3.8	593	1	Y40A_RHITSN	P95586	rhizobium s
23	104.5	3.8	516	1	ER13_ARATH	Q65403	arabidopsis
24	103	3.8	472	1	PCNB_ECOLI	P13685	escherichia
25	103	3.8	510	1	C4DK_DRONE	Q3w011	tyrosophila
26	101.5	3.7	372	1	FTZL_PYRHO	Q57776	pyrococcus
27	101.5	3.7	556	1	FTHS_STRMU	Q59925	streptococc
28	100.5	3.7	366	1	FTZL_PYRWO	Q52630	pyrococcus
29	100.5	3.7	497	1	TRXB_HUMAN	Q16881	homo sapien
30	100	3.6	1071	1	CARB_BACSU	P25994	bacillus su
31	100	3.6	1202	1	ALAA_ARATH	Q91183	arabidopsis
32	99.5	3.6	308	1	R34K_CLOPA	P23160	clostridium
33	99.5	3.6	1064	1	CARB_LACIC	Q32771	lactococcus

Query Match 69.6%; Score 1909.5; DB 1; Length 526;
 Best Local Similarity 68.5%; Pred. No. 5.5e-139;
 Matches 369; Conservative 68; Mismatches 78; Indels 23; Gaps 6;

QY 115 CGPAGLSLAALAAKGLKVLGIGLDPFTNNYGVWDEDFKDLGLERCIEHAKDTIIVYLD 174
 Db 116 CGPAGLALAAESAKGLKVLGIGLDPFTNNYGVWDEDFKDLGLERCIEHAKDTIIVYLD 175
 QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGSVYLDKVERITEAGDGHSLVVCENEIFI 234
 Db 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGSVYLDKVERITEAGDGHSLVVCENEIFI 235
 QY 235 PCRLATVASGAASGLILEYEVGGPRVCQVAYGVEVEVNNPYDPNLMVMDYRDYMOOK 294
 Db 236 PCRLATVASGAASGLILOEYEVGGPRVCQVAYGVEVEVNNPYDPNLMVMDYRDYMOOK 295
 QY 295 LQSESEYPTFLYVMPSPRLFFETCLASKDAMPDLKRLKMSRLKTLIGIQVTKYVE 354
 Db 296 VRSLEAEYPTFLYVMPSPRLFFETCLASKDAMPDLKRLKMSRLKTLIGIQVTKYVE 355
 QY 355 EENSYPVGGSLPNTQKNLAFGAASAMVHPATGYSVVRSLSEAPKASYATKILKQDNS 414
 Db 356 EENSYPVGGSLPNTQKNLAFGAASAMVHPATGYSVVRSLSEAPKASYATKILKQDNS 415
 QY 415 AYVVGSSAVNISQAWSSLPKRRQRRAFFLGLGLIVQLDIEATRTFFFRFLPT 474
 Db 416 KQINS-----NISQAWDTLMPERRKRRQRAFFLGLGLIVQLDIEATRTFFFRFLPT 475
 QY 475 WMWNGFLGSSLSFDLVFMSYMFVLPAPNSMRSLVRHLLSDPBGAVVWVAYLE 528
 Db 470 WMWNGFLGSSLSFDLVFMSYMFVLPAPNSMRSLVRHLLSDPBGAVVWVAYLE 523

RESULT 2

LCYB_LYCES STANDARD; PRT; 526 AA.
 AC 065837;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lycopen epsilon cyclase, chloroplast precursor (EC 1.14.-.-).
 GN CRTL-E-1.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=Leaf;
 RA Ronen G., Cohen M., Zamir D., Hirschberg J.;
 RT "Regulation of expression of the gene for lycopen epsilon cyclase
 during fruit ripening of tomato.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: CATALYZES THE SINGLE CYCLIZATION REACTION WHICH CONVERTS
 CC LYCOPENE TO DELTA-CAROTENE AND NEUROSPORENE TO ALPHA-ZEACAROTENE.
 CC -|- SUBCELLULAR LOCATION: Chloroplast.
 CC -|- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Y14387; CAA74745.1;
 KW Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 KW Transit peptide.
 FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
 FT CHAIN ? 526 LYCOPENE EPSILON CYCLASE.
 FT NP_BIND 108 136 NAD (POTENTIAL).
 SQ SEQUENCE 526 AA; 58886 MW; 37357C3869DBDCAF CRC64;

RESULT 3

LCYB_CAPAN STANDARD; PRT; 498 AA.
 AC Q43415;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lycopen beta cyclase, chloroplast precursor (EC 1.14.-.-).
 GN LCY1 OR CRTL.
 OS Capsicum annuum (Bell pepper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
 OX NCBI_TaxID=4072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Lamuyo; TISSUE=Fruit;
 RX MEDLINE=96045549; PubMed=7550379;
 RA Huguency P., Badillo A., Chen H.C., Klein A., Hirschberg J.,
 RA Camara B., Kuntz M.;
 RT "Metabolism of cyclic carotenoids: a model for the alteration of this
 RT biosynthetic pathway in Capsicum annum chromoplasts.";
 RL Plant J. 8:417-424(1995).
 CC -|- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS
 CC LYCOPENE TO BETA-CAROTENE AND NEUROSPORENE TO BETA-ZEACAROTENE.
 CC -|- PATHWAY: Carotenoid biosynthesis.
 CC -|- SUBCELLULAR LOCATION: Chloroplast; chromoplast.
 CC -|- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X86221; CAA60119.1; -
 InterPro: IPR000205; NAD_binding.
 Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 Transit peptide.
 TRANSIT 1 79 CHLOROPLAST (POTENTIAL).
 CHAIN 80 498 LYCOPENE BETA CYCLASE.
 NP_BIND 84 112 NAD (POTENTIAL).
 SEQUENCE 498 AA; 55610 MW; 177180CD5745F64F CRC64;
 Query Match 27.7%; Score 761; DB 1; Length 498;
 Best Local Similarity 35.1%; Pred. No. 7.3e-51;
 Matches 174; Conservative 100; Mismatches 184; Indels 38; Gaps 12;
 QY 43 FQVRADGGSGRS---SVAYKEGVDEDFIKAGGSELL-FVQMQTKSMKQAKLADK 97
 DB 17 FGKVSASFSSVKSGFGAKKFCGLGSRVSVCKASSALLELVPEKKENLDFELPMYD- 75
 QY 98 LPPPIFGESVMDLVVIGCGPAGLSAAEAAGLKVGLIGPD--LPFTNNYGVWEDEPKD 155
 DB 76 ---PSKGVVVDLAVVGGPAGLAAQVSEAGLSVCSIDPNPKLIWPNNGYVWVDEFEA 131
 QY 156 LGLERCIEHAKRDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVER 215
 DB 132 MDLLDCLDATNSGAAYVYDDTKDLNRPYGRVNRKQLSKMMQKCIILNGVFHQAQVTK 191
 QY 216 ITEAGDGHSLVVCNEIFIPCRLATVAGSAAGSKLLEYEVG-GPRVCVQATYGEVEVEN 274
 DB 192 VIHE-ESKSMCLCNDGITIATVVDATG-PSRLVQVDKPNPG--YQVAYGILAEVEE 247
 QY 275 NPYDNLAVFYDYRD-YMQOKLOCSE--EYPTFLYVMPMSTRLEFETCLASDAMPF 331
 DB 248 HFDVKNVFMWDRSLKNNVLEKERNRIPTFLYAMPFSSNRIFLEETSLVAPRGGM 307
 QY 332 DLLKRLMSRLKTLGIQVTKVVEEWSYIPVGGSLPNTQKLNAGFAASWVHPATGYV 391
 DB 308 DDIERMVARLSHLGIKVKYSIEDEHCVIPMGPLPVPQVVGIGGTAGVHPSTGYV 367
 QY 332 VRSLEAPYASVIAKILKQDNASVYVQSGSAVINISQAWNSLMPKPKRORAFFLGL 451
 DB 368 ARTLAAPVPAVNAIIQYLSERS-----HSGDELSAAVWKLWPIERRRQREFFCFGM 420
 QY 452 ELIVOLDTEATRTFTFRLPTWMMGFLGSLSSFDLVFSMYMFLVAPNSRMS--- 508
 DB 421 DILLKLDLPATRRFFDAFDLEPRYWHGFLSLRLFLPELVFGLSLFASHASNTSRLEIMT 480
 QY 509 -----LVRLHLS 516
 DB 481 KGTPLVHMINLLQD 496

RESULT 4

LCVB_LYCES
 ID LCVB_LYCES STANDARD; PRT; 500 AA.
 AC Q43503;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lycopen beta cyclase, chloroplast precursor (EC 1.14.-.-).
 GN LCY1 OR CRT1-1.
 OS Lycopersicon esculentum (Tomato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 RN NCBI_TaxID=4081;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=Leaf;
 RX MEDLINE=96434545; PubMed=8837512;

RA Cunningham F.X. Jr., Pogson B., Sun Z., McDonald K.A., Dellapenna D.,
 Gantt E.;
 RT "Functional analysis of the beta and epsilon lycopene cyclase enzymes
 of Arabidopsis reveals a mechanism for control of cyclic carotenoid
 formation.";
 RL Plant Cell 8:1613-1626(1996).
 CC -!- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS
 LYCOPENE TO BETA-CAROTENE AND NEUROSPORENE TO BETA-ZEACAROTENE.
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL: X86452; CAA60170.1; -
 InterPro: IPR000205; NAD_binding.
 Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 Transit peptide.
 TRANSIT 1 81 CHLOROPLAST (POTENTIAL).
 CHAIN 82 500 LYCOPENE BETA CYCLASE.
 NP_BIND 86 114 NAD (POTENTIAL).
 SEQUENCE 500 AA; 56180 MW; CF42F7D4684C04DD CRC64;
 Query Match 27.3%; Score 749.5; DB 1; Length 500;
 Best Local Similarity 36.6%; Pred. No. 5.6e-50;
 Matches 169; Conservative 92; Mismatches 168; Indels 33; Gaps 11;
 QY 73 GSELLFVQMQTKSMEKAKLADKLPPT-PFGESVMDLVVIGCGPAGLSAAEAAGL 131
 DB 52 GSSALLELVPEP-----KKENLDFELPMYDPSKGVVVDLAVVGGPAGLAAQVSEAGL 107
 QY 132 KYGLIGPD--LPFTNNYGVWEDEKDLGERCIEHAKRDTIVYLDNDAPVLIGRAYGRVS 189
 DB 108 SVCSIDPNKLIWPNNGYVWVDEFEAMDLLDCLDATWSGAAYVDDTAKDLHPYGRVN 167
 QY 190 RHLLEELKRCVSGSVYLSKVERITEAGDGHSLVVCNEIFIPCRLATVAGSAAGSK 249
 DB 168 RKQLSKMMQKCIILNGVFHQAQVVIHE-ESKSMCLCNDGITIATVVDATG-FSRS 225
 QY 250 LLEYEVG-GPRVCVQATYGEVEVENNPDNLVFMFYDYRD-YMQOKLOCSE--EYPTF 305
 DB 226 LVOYDYPYNG--YQVAYGILAEVEEHPFDVKNVFMWDRSLKNNVLEKERNRIPTF 283
 QY 306 LYVMPMSTRLEFETCLASDAMPFLLKRLMSRLKTLGIQVTKVVEEWSYIPVGG 365
 DB 284 LYAMPFSSNRIFLEETSLVAPRGILRIDDIOERMVARLNHLGKVKSEDEHCLIPMG 343
 QY 366 LPNTEOKNLAFGAASWVHPATGYSVVRSLSSEAPKASVATAKILKQDNASVYVGG 425
 DB 344 LPVLPQVVGIGGTAGVHPSTGYMVARTLAAAPVPAVNAIIQYLSERS-----HSGN 396
 QY 426 NISMQAWSLMPKPKRORAFFLGLFLVQLDIEATRTFTFRLPTWMMGFLGSSL 485
 DB 397 ELSTAVKWLWPIERRRQREFFCFGMDDLLKLDLPATRRFFDAFDLEPRYWHGFLSSRL 456
 QY 486 SSFDFLVFSMYMFLVAPNSMR-----MSLVRLHLS 516
 DB 457 FLPELVFGLSLFASHASNTSRFEIMTKGTVPVLMNINLLQD 498

RESULT 5
 LCVB_ARATH
 ID LCVB_ARATH STANDARD; PRT; 501 AA.
 AC Q38933; Q39145;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

[illegible]

CCS_CAPAN STANDARD; PRT; 498 AA.

AC Q42435; Q39470;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Capsanthin/capsorubin synthase, chloroplast precursor.

GN CCS.

OS Capsicum annuum (Bell pepper).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.

OX NCBI_TaxID=4072;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Yolo Wonder;

RX MEDLINE=94197697; PubMed=8147854;

RA Derriere J., Bouvier F., Steppuhn J., Klein A., Camara B., Kuntz M.;

RT "Structure and expression of two plant genes encoding chloroplast-

RT specific proteins: occurrence of partially spliced transcripts.";

RL Biochem. Biophys. Res. Commun. 199;1144-1150(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Lamuyo;

RX MEDLINE=95004653; PubMed=7920703;

RA Bouvier F., Huquaney P., d'Harlinque A., Kuntz M., Camara B.;

RT "Xanthophyll biosynthesis in chloroplasts: isolation and molecular

RT cloning of an enzyme catalyzing the conversion of 5,6-epoxycarotenoid

RT into ketocarotenoid.";

RL Plant J. 6:45-54(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Yolo Wonder;

RX MEDLINE=95179811; PubMed=7874747;

RA Houl G., Schantz M.L., Meyer B., Pozueta-Romero J., Schantz R.;

RT "A chloroplast-specific protein in Capsicum annum: characterization

RT and expression of the corresponding gene.";

RL Curr. Genet. 26:524-527(1994).

CC -!- FUNCTION: CATALYZES THE CONVERSION OF THE UBIQUITOUS 5,6-

CC EPOXYCAROTENOID, ANTHEXANTHIN AND VIOLAXANTHIN, INTO CAPSANTHIN

CC AND CAPSORUBIN, RESPECTIVELY.

CC -!- PATHWAY: Carotenoid biosynthesis.

CC -!- SUBUNIT: MONOMER.

CC -!- SUBCELLULAR LOCATION: Chloroplast; chromoplast.

CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X77289; CA54495.1; -

DR EMBL; X76165; CA53759.1; -

DR EMBL; X78030; CA54961.1; -

DR InterPro: IPR000205; NAD_binding.

KW Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;

KW Transit peptide.

FT TRANSIT 1 ?

FT CHAIN 1 ?

FT NP_BIND 84 498

FT CAPSANTHIN/CAPSORUBIN SYNTHASE.

FT NAD (POTENTIAL).

FT A -> R (IN REF. 3).

FT CONFLICT 316 316

FT CONFLICT 378 380

FT AEA -> LRP (IN REF. 3).

FT CONFLICT 405 408

FT PSD -> LRP (IN REF. 3).

FT CONFLICT 458 497

FT ELAVSLYLFGRASNLARDIVKTCVPLVKLGNLAIES

FT -> RTCCTQFVFWTC (IN REF. 3).

FT SEQUENCE 498 AA; 56658 MW; 4FB25F676A022A98 CRC64;

Query Match 25.7%; Score 705.5; DB 1; Length 498;

Best Local Similarity 38.6%; Pred. No. 1.3e-46;

Matches 158; Conservative 79; Mismatches 153; Indels 19; Gaps 9;

QY 109 DLVVIGCGPAGLSLAEEAAKGLKVLGIGDLPFT---NNYGVWEDEFKDLGLERCIEHA 165

DB 83 DVIILGTGAGRLAEQVSKYGIKVCVDPSP-PLSMWPNNGYVWVDEFEKLGLEDCLDHK 141

QY 166 WKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLARCVEGYSYLDKSKVERITEAGDGHSL 225

DB 142 WPVSCVHISDHKTKYLDPRYGRVSRKKLKLKLLNSCVENRYKFKAKYLVKHE-EFES 200

QY 226 VVCNEIFIPCLATFASGAAGSKLLEVEVGPR-VCVQTAYGVEVEVNNPYDNLWVF 284

DB 201 IVDDGRKISGLVDASGYAS-DFIEYD--KPRHGYQVAHGILAEVDNIPFDLKKML 257

QY 285 MDYRDY---MQOKLQCSEEEYPTFLYVMPSPTRLFFETCTCLASKDAMPFLLKRLMSR 341

DB 258 MDWRDHLGNEPYLRVKNTRKPTFLYAMPFDRLNLFLEETSILVSRPMLSYMEVKRMVAR 317

QY 342 LKTLGIQVTKVYEENSYIPVGGSLPNTQKNLAFGAASVMVHPATGYSVVSRSLEAPKY 401

DB 318 LRHLGIKVRSLVEEEKCVITMGSGPLPRIPQNVMAIGTSGIVHPSSGYMARSALAPVL 377

QY 402 ASVIATKILKQDNSAYVVGSSAVNISQAWSSLWPKRKRORAFFLGLLELIYOLDIEA 461

DB 378 AEAVESL---GSTRMIRGS---QLYRWVWNLWPSDRRVRECYCFGMETLKLKLDLEG 430

QY 462 TRTFFRTFFRLPTWMMWGLGSSSFDLVLFMSYMFVLPAPNSMRSLV 510

DB 431 TRRLDFAFFDPRYWHGFLSSRLSVKELAVLSLYLFCHASNARLDIV 479

RESULT 9

LCYB_NARPS STANDARD; PRT; 503 AA.

AC Q40424;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lycopen beta cyclase, chloroplast precursor (EC 1.14.-.-).

GN LCY1 OR LYC.

OS Narcissus pseudonarcissus (Daffodil).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;

OC Narcissus.

OX NCBI_TaxID=39639;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Paracolla;

RA Al-Babili S., Hobeika E., Beyer P.;

RT "A cDNA encoding lycopene cyclase from Narcissus pseudonarcissus L.";

RL (in) Plant Gene Register PGR96-107.

RN [2]

RP SUBCELLULAR LOCATION.

RX MEDLINE=97433278; PubMed=9288918;

RA Bonk M., Hoffmann B., von Lintig J., Schledz M., Al-Babili S.,

RA Hobeika E., Kleinig H., Beyer P.;

RT "Chloroplast import of four carotenoid biosynthetic enzymes in vitro

RT reveals differential fates prior to membrane binding and oligomeric

RT assembly.";

RL Eur. J. Biochem. 247:942-950(1997).

CC -!- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS

CC LYCOPENE TO BETA-CAROTENE AND NEUROSPORENE TO BETA-ZEACAROTENE.

CC -!- PATHWAY: Carotenoid biosynthesis.

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST; CHROMOPLAST. EXISTS AS AN

CC INACTIVE SOLUBLE FORM AND AN ACTIVE MEMBRANE-BOUND FORM

CC (PROBABLE).

CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

or send an email to license@sib-sib.ch.

```

CC -----
CC EMBL: X98796; CAAG7331.1;
CC Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
CC TRANSIT PEPTIDE; Membrane.
KW TRANSIT 1 85 CHLOROPLAST (POTENTIAL).
FT CHAIN 86 503 LYCOPENE BETA CYCLASE.
FT NP_BIND 90 117 NAD (POTENTIAL).
SQ SEQUENCE 503 AA; 56910 MW; 3FD1E355EF184D98 CRC64;

Query Match 25.6%; Score 701.5; DB 1; Length 503;
Best Local Similarity 36.4%; Pred. No. 2.7e-46;
Matches 159; Conservative 87; Mismatches 160; Indels 31; Gaps 11;

QY 93 KLADKLPP-----PFGESVMDLVVIGGPGAGLSLAAEAAKLGKVLGIGPD 139
DB 60 QLLDLVPELKEHEFDLPYDPSKALTLDLAVVGGPLARSCSTSLGG-GLSVSIDPN 118

QY 140 --LPFTNNGVWEDEPKDLGLERCIEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEEL 197
DB 119 PKLIWPNNGVGVWEDEPKDLGLERCIEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEEL 178

QY 198 LKRCVSGVSYLDSKVERITEAGDGHSLVCENEIIPICRLATVAGSAAGSKLLEVEVG- 256
DB 179 MKKCSVNGVRFHQATVVKAMHE-EKSYLTCSGVTIDARVVLDTAG-FSRCLVQYDKPY 236

QY 257 GPRVCVQATAYGVEVEENPNPDMVMDYRD-YMOQKLOCSEE-EYPTFLYVMPMS 313
DB 237 NPG--YQVAYGILAEVEEHPFDVKMVDWDRDLSHNGKALNERNAKPTFLYAMPFSS 294

QY 314 TRLEFEETCLASKDAMPFDLLKRLKMSRLTKLIGIQVTKYVEEWSVIPVGGSLPNTQKN 373
DB 295 NRIFLEETSILVAPGLKMDIQRVAVRNHLGIRIKSTIEEDERCVPIMGGLPVPQPV 354

QY 374 LAFGAASVMVHPATGYSVVRSLEAPKYAVIAKILKQNSAVVSGQSSAVNISMQAWS 433
DB 355 VGIGTAGVMVHPSTGTWARTAAAVANSIVQYLVSDSG---LSGN----DLSADVWK 407

QY 434 SLWPKRRQRRAFFLGLLEIVQLDIEATRTFTFRPLPTMWMWGLGSSLSFDLVLF 493
DB 408 DLWPIERRRQREFFCFGMDILLKLDLGTRFFDAFFDLEPRYWHGFLSRFLPELVFP 467

QY 494 SNYMFVLAPNSMRSLV 510
DB 468 GLSLFASHASNTCKLEIM 484

RESULT 10
LCYB_SYN7
ID LCYB_SYN7 STANDARD; PRT; 411 AA.
AC Q55276;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lycopen beta cyclase (EC 1.14.-.-).
GN CRTL OR LCY.
OS Synchococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95003701; PubMed=7919981;
RA Cunningham F.X. Jr., Sun Z., Chamovitz D., Hirschberg J., Gantt E.;
RT "Molecular structure and enzymatic function of lycopene cyclase from
RT the cyanobacterium Synchococcus sp strain PCC7942."
RL Plant Cell 6:1107-1121(1994).
CC -!- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS
CC LYCOPENE TO BETA-CAROTENE AND NEUROSPORINE TO BETA-ZEACAROTENE.
CC -!- ENZYME REGULATION: INHIBITED BY THE BLEACHING HERBICIDE 2-(4-
CC METHYLPHENOXY)TRILETHYLAMINE HYDROCHLORIDE (NPTA).
CC -!- PATHWAY: Carotenoid biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X74599; CAAS2677.1;
DR InterPro: IPR00205; NAD_binding.
KW Oxidoreductase; NAD; Carotenoid biosynthesis.
FT NP_BIND 4 32 NAD (POTENTIAL).
SQ SEQUENCE 411 AA; 46085 MW; C46CC5B2E85E7AC2 CRC64;

Query Match 21.9%; Score 602; DB 1; Length 411;
Best Local Similarity 35.5%; Pred. No. 8.9e-39;
Matches 150; Conservative 61; Mismatches 165; Indels 46; Gaps 10;

QY 107 VMDLVVIGGPGAGLSLAAEAAKLGKLV-GL--IGPDLPTFTNNGVWEDEPKDLGLERCIE 163
DB 1 MEDALVIGSGPGAGLAAELAAQRLKVOGLSPVDPFHPWENTYGIWGPDLDSGLGLEHFG 60

QY 164 HAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEELKRCVSGVSYLDSKVERITEAGDGH 223
DB 61 HRWSNCVSYF-GEAPVQHQNYGLFDRAQLQOHWRLOCEOGGLQWLGAAL--AHDSh 117

QY 224 SLVVCENEIIPICRLATVAGSAAGSKLLEVEVGPR-----VCVOTAYGVEVEV 272
DB 118 HSCV-----TTAAGQELQARLVYDTTGHQAAFTQPHSDAIAQAAVIGIQGF 165

QY 273 ENNPYDNLVMDYR-DYMOQKLOCSEEEY---PTFLYVMPMSPTLFFETCLASKDA 328
DB 166 SQPIEPHQFVLMYRSDHL-----SPEERQLPPTFLYAMDLDNDVVFVEETSIAACPA 219

QY 329 MPFDLLKRLKMSRLTKLIGIQVTKYVEEWSVIPVGGSLPNTQKNLAFGAASVMVHPATG 388
DB 220 IPYDLRKQRLQRLATRGVTQVIOHEEYCLFPMNLPLDPLTQSVGVGGGAASVMVHPASG 279

QY 389 YSVVRSLEAPKYAVIAKILKQNSAVVSGQSSAVNISMQAWSLSLWPKRRQRRAFFL 448
DB 280 YMVGALLRRAPDLANAAGLNASSL-----TTAELATQAWRGLWPTKIRKHYIQ 332

QY 449 FGLELIVQLDIEATRTFTFRPLPTMWMWGLGSSLSFDLVLFMSYMFVLAPNSMRMS 508
DB 333 FGLEKLMRFSEAQLNHHFQTFGLPKBQWYGFITNTLSLPELIQAMLRLEFAQAPNDVRWG 392

QY 509 LV 510
DB 393 LM 394

RESULT 11
Y801_DEIRA
ID Y801_DEIRA STANDARD; PRT; 410 AA.
AC Q9RW68;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative carotenoid cyclase DR0801.
GN DR0801.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

```

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "genome sequence of the radioresistant bacterium Deinococcus
 RL radiodurans R1.";
 RT Science 286:1571-1577(1999).
 CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF001934; AAF10377.1; -
 DR TIGR; DR0801; -
 DR InterPro; IPR003042; Rng_mnoxygenase.
 DR PRINTS; PR00420; RNMNOXGNASE.
 KW Hypothetical protein; Oxidoreductase; NAD; Carotenoid biosynthesis;
 KW Complete proteome.
 FT NP_BIND 11 39 NAD (POTENTIAL).
 SQ SEQUENCE 410 AA; 43161 MW; E1B3162F10F9E6AF CRC64;
 Query Match 15.7%; Score 429.5; DB 1; Length 410;
 Best Local Similarity 31.0%; Pred. No. 1.5e-25;
 Matches 135; Conservative 53; Mismatches 204; Indels 43; Gaps 13;
 QY 102 PFGESVMDLVITGCGPAGLSLAAEAKLGLKVLIG--PDLPFTNNYGVWEDEFKDLG-- 157
 DB 3 PFPASSDVLVITGGPGSGTALSAAELAGLDVQQLAPHPPPFPATYGAN-----LGLD 56
 QY 158 ---LERCIEHAKDVIYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVE--SGVSYLDSK 212
 DB 57 PFWAGCAEQVWTDVRAY-TGQPSTSLGQPYA-----LLDNAALLTLRLGLADMTWVEGA 110
 QY 213 VERITEAGDGHSLVVCNEIEIFPCRLATVAGSAAGSKLLEYEVGGPRVCVOTAYGVEVEV 272
 DB 111 ALHAERSGAGWTVYAGGERW-QTFLVVDAGS--HGALVSPYRFPFGAALQATAYGVWARF 167
 QY 273 ENNPYPDNLVMDYRDYMQKLCQSEEEYPTFLVYMPSPTRLFTECLASKDAMPFD 332
 DB 168 RRPVPTGSMWMDYRT-PAPELARGE---ATFLYAMHLGGDRVFEVETSLIARPAITRA 223
 QY 333 LKRRKLSRLKTLGIQVTKVVEEWSYIPVGGSLPNTPEOKNLAFGAASWHPATGYSV 392
 DB 224 ELRRELLARLSAQTPPHATSEEWAFPMNAQAP-APGGVLAAYGAAAGRVHPVSGFOVA 282
 QY 393 RSLSEAPKYASVIAKILKQDNSAYVVSQSSAVNISQWSSSLWPKRKRORAFELFGL 452
 DB 283 GALSADAPGVATAIATLCCQKDA-----AAGWAALWSPERRAAREVHLLGVG 330
 QY 453 LIVOLDIEATRTFTFTRLPVWVGLASSLSDFDLVFSMTVMFLVAPNSMRLVRH 512
 DB 331 ALLGLERAEHPFGFTFGDLPREQWAFRLHPDPTDAGTLARTMLRVFAQTGGRVRLPLARA 390
 QY 513 LLSDP--SGAVMVA 525
 DB 391 ALAQPAASGRALAAA 405
 RESULT 12
 ERIL_BRANA STANDARD; PRT; 506 AA.
 AC 065727;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Squalene monoxygenase 1.1 (EC 1.14.99.7) (Squalene epoxidase 1.1) (SE
 DE 1.1).
 GN SQP1.1.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Westar; TISSUE=Shoot;
 RX MEDLINE=99277589; PubMed=10350086;
 RA Schaffer U.A.; Reed D.W.; Hunter D.G.; Yao K.; Weninger A.M.;
 RA Tsang E.W.T.; Reaney M.J.T.; Mackenzie S.L.; Coveillo P.S.;
 RT "An example of intron junctional sliding in the gene families encoding
 RT squalene monoxygenase homologues in Arabidopsis thaliana and
 RT Brassica napus.";
 RT Plant Mol. Biol. 39:721-728(1999).
 CC -!- FUNCTION: CATALYZES THE FIRST OXYGENATION STEP IN STEROL
 CC BIOSYNTHESIS AND IS SUGGESTED TO BE ONE OF THE RATE-LIMITING
 CC ENZYMES IN THIS PATHWAY.
 CC -!- CATALYTIC ACTIVITY: Squalene + AH(2) + O(2) -> (S)-squalene-2,3-
 CC epoxide + A + H(2)O.
 CC -!- COFACTOR: FAD.
 CC -!- SIMILARITY: BELONGS TO THE SQUALENE MONOOXYGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ005931; CAA06773.1; -
 DR InterPro; IPR000733; Flav_monoxygenase.
 DR InterPro; IPR000205; NAD_binding.
 DR Pfam; PF01360; Monoxygenase; 1
 KW Oxidoreductase; Flavoprotein; FAD; Transmembrane; Multigene family.
 FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 47 67 POTENTIAL.
 FT TRANSMEM 447 467 POTENTIAL.
 FT NP_BIND 50 77 FAD (ADP PART) (POTENTIAL).
 SQ SEQUENCE 506 AA; 55571 MW; 32C0301F1A66CD13 CRC64;
 Query Match 5.0%; Score 137.5; DB 1; Length 506;
 Best Local Similarity 24.3%; Pred. No. 0.0053;
 Matches 108; Conservative 52; Mismatches 145; Indels 139; Gaps 25;
 QY 89 EKAQKADLKLPPIPFGESVMDLVITGCGPAGLSLAAEAKLGLKVLIGDPL--PFTNNY 146
 DB 30 KKVAKLPDAATEYR-RGDADVIIVGAGVGSALAYALAKDGRVHVIERDMREPVR--- 85
 QY 147 GVWEDEF-----KDLGERCIE-----HAWKDTIYVL-----DNDAPV-LI 181
 DB 86 --MMGEFMPQGGRLLLSKLGLLEDCLEGIDQIATGLAVYKDGKALVSFFEDNDPVEPT 143
 QY 182 GRAY--GRVSRHL-----LHEELLKRCVE-----SGVSYLDSKVERTEAGDGH 223
 DB 144 GRAFYNGRFVQRLRQKASSLPTVQLEBGTVKSLEEKGVTKGVTKYNS-----AGE-- 194
 QY 224 SLVVCNEIPIPCRLATVAGSAAS-----GKLLLEYEVGGPRVCVOTAYGVEVE 273
 DB 195 -----ETTAPAP--LTVVCDGCVSNLRRSVNDNNAEIVSYQVG-----YVSKNCQLE-- 239
 QY 274 NNYPDNLVMDYRDYMQKLCQSEEEYPTFLVYMPSPTRLFTECLASKDAM----- 329
 DB 240 -----DPEKLLKMSKPSFTMLYQIISSTDVRCVMEIFPGNIPISNGEMAYILKNTMAPQV 295
 QY 330 PFDLLKRLKLSRLKTLGIQV---TKVYEEWSYIPVGGSLPNTPEOKNLAFGAASWHP 385
 DB 296 PPEL--RKIFLKGIDSGAQIKAMPTRKMEATLS-----EKGVIVLGDAPFNNRHP 343
 QY 386 --ATGYSVY-----RSLSEAPKYASVIAKILKQDNSAYVY--SQGSSAVNIS 428
 DB 344 ATASGMVMVYLSLILRLRLLOPLRLNLSKANKVSEVI-----KSFYVIRKPKMSATVNTL 396
 QY 429 MQAWSSSLW-----PKRKRQRAF 446

DB	299	IRDYAI---DRARDRELRLNRMFLPGCAPDRRYTLQRFYRMPHGLIERFAGRLSVA	355
QY	489	D 489	
DB	356	D 356	
RESULT 14			
CRTY_PANAN	STANDARD;	PRT;	382 AA.
ID	CRTY_PANAN		
AC	P21687;		
DT	01-MAY-1991 (Rel. 18, Created)		
DT	01-MAY-1991 (Rel. 18, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Lycopene cyclase.		
GN	CRTY.		
OS	Pantoea ananas (Erwinia uredovora).		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Pantoea.		
OX	NCBI_TaxID=553;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=20D3;		
RX	MEDLINE=91072214; PubMed=2254247;		
RX	Misawa N.; Nakagawa M.; Kobayashi K.; Yamano S.; Izawa Y.,		
RT	Nakamura K.; Harashima K.;		
RT	"Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway		
RT	by functional analysis of gene products expressed in Escherichia		
RT	coli.";		
RL	J. Bacteriol. 172:6704-6712(1990).		
CC	-I- FUNCTION: CATALYSES THE CYCLIZATION REACTION WHICH CONVERTS		
CC	LYCOPENE TO BETA-CAROTENE.		
CC	-I- PATHWAY: Carotenoid biosynthesis.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)		
CC	or send an email to license@isb-sib.ch .		
CC	-----		
DR	EMBL; D90087; BAA14126.1; -		
DR	PIR; C37802; C37802.		
KW	Carotenoid biosynthesis.		
SEQ	SEQUENCE 382 AA; 43047 MW; 62A94222A9EBD45 CRC64;		
Query Match 4.6%; Score 125.5; DB 1; Length 382;			
Best Local Similarity 20.0%; Pred. No. 0.03;			
Matches 79; Conservative 74; Mismatches 185; Indels 57; Gaps 15;			
QY	109	DIYVIGGPGAG--LSLAAEAAGLKVGLIGDLPFTNNYGVWEDEFKDL-----GLER	160
DB	6	DLILVGLAGLIALRLQOQDPMDRLILDA-APQAGNHTWSFHDDLTESQHRWIAP	64
QY	161	CIEHAWKDTTVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVESGV-S-YLDSKYVERITEA	219
DB	65	LVVHHWPDYQVRPPTRRKLSGVFCITSQRF--AEVLQR--QGFPHLWMDTAVAEV---117	
QY	220	GDGSLVSCVNEIFIPORLATVAGSAAGSKLLEYEGGPRVCVQTATGVEVEVENNPDP	279
DB	118	-NAESVRLKKGQV-IGARAVIDGRGYAANSALS-----VGFOAFIGQEWRL-SHPHGL	167
QY	280	NLMVFMDYRDYMOOKLQCSSEETFLYVMPMSPTRLFEEETCLASKDAMPDILLKRLM	339
DB	168	SSPLIM-----AVDQONGRFYFVPLSPLETHIVIDNATLDPECARONIC	219
QY	340	SLRKTGLIGQTKVVEEWSYIPV---GCSLPNTBQKNLA-FGAASVHPHATGYSVVRSL	395
DB	220	DYAAQQGWOQLTLREBQALPITLSNADAFWQORPLACSLRAGLPHPTTGSPLAV	279
QY	396	SEAPKYASVIATILKDNASVYVYVSGSSAVNISMOAWSSLMPPKRRQRAPFF-LFGLELI	454

Db 280 AVADR-----LSALDVTSASIIHATHFARERWQQGFFRLNRMFL 322
QY 455 VOLDIEATRTFTFRRLPTWMWCFGLGSSLSSED 489
Db 323 LAGPADSRWRYMQRFGPLPDLIARFYAGKLTLD 357

RESULT 15
CRTV_ERWHE
ID CRTV_ERWHE STANDARD; PRT; 386 AA.
AC Q01331;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lycopene cyclase.
GN CRTV.
OS Erwinia herbicola.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EH010;
RX MEDLINE=93138098; PubMed=8422926;
RA Hundle B.S., O'Brien D.A., Beyer P., Kleinig H., Hearst J.E.;
RT "In vitro expression and activity of lycopene cyclase and
beta-carotene hydroxylase from Erwinia herbicola.";
RL FEBS Lett. 315:329-334(1993).
CC -1- FUNCTION: CATALYSES THE CYCLIZATION REACTION WHICH CONVERTS
LYCOPENE TO BETA-CAROTENE.
CC -1- PATHWAY: Carotenoid biosynthesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M87280; AAA64980.1; -
KW Carotenoid biosynthesis.
SQ SEQUENCE 386 AA; 43341 MW; F4A40563BFCFA980 CRC64;

Query Match 4.68; Score 125; DB 1; Length 386;
Best Local Similarity 20.3%; Pred. No. 0.033;
Matches 86; Conservative 62; Mismatches 163; Indels 112; Gaps 18;

QY 109 DLVVIGCPA-GL---SLAAEAAKGLKVLGPDLPFTNNYGVWEDEF---KDLGLERC 161
Db 3 DLIIVGGLANGLIANLRLQRYPOLNLLIEAGEQPGCNHTWSEFDDLPQGHAWLAPL 62
QY 162 IEHAKDTIYVLDNDAPVLIGRAYGRVSRHLHEELKRCVSGSVYSDSKVERITEA-- 219
Db 63 VAHAWPGYEQFPD-----LRRRLARG--YYSITSERFAEALH 98
QY 220 ---GDGHSVLVCNEIFPCRLATV-----ASGAA--SGKLLLEVGGPRVCVQTAYG 267
Db 99 QALGE-----NWLNCVSEVLPSNVRNGLANGALLAGAVIDGRGVTVASSAMOTGYQ 149
QY 268 VEVEVE---NNPYDPLMVDYDMQKLOCSEEEYPTFLYVMPSPTRLFEEETCLA 324
Db 150 LFLGQQWRLTQPHGLVPIILMDATVAQQQYR-----FVYTLPSADTLIEDTRYA 201
QY 325 SKDAMPDLKRLKMLSKLTIGTQVTKVVEEWSYIPV--GGSL-----PNTQKNLA 375
Db 202 NVPQRDNDALRQVTDYAHKSGWQLAQERETGCLPITLAGDIQALWADAPGVPRS--- 258
QY 376 FGAAASWVHPATGYS-----VVRSLSEAPKIASVTAKILKQNSAYVVGSGSSAVNIS 429
Db 259 -GMRAGLFHTTGYSLPLAVALADAIADSPRLGSV-----PLYQLTR 299
QY 430 QAWSSLPKE---RRQRRAFFLFGLELIVQLDIEATRTFTFRRLPTWMWCFGLGSSLS 486

Search completed: May 21, 2003, 22:09:40
Job time : 27 secs

Db 300 QFAERHWRQGFRRLLNMLFLAGRE-----ENRNRVMQRFYGLPEPTVVERFYAGRLS 352
QY 487 SFD 489
Db 353 LFD 355

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: May 21, 2003, 22:05:08 ; Search time 72 Seconds
(without alignments)
1513.874 Million cell updates/sec

Title: US-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRLNLISSCPVWTFGT.....VRHLLSPGAVMVRAYLER 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 205047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP-archaea.*
- 2: SP-bacteria.*
- 3: SP-fungi.*
- 4: SP-human.*
- 5: SP-invertebrate.*
- 6: SP-mammal.*
- 7: SP-mhc.*
- 8: SP-organalle.*
- 9: SP-phage.*
- 10: SP-plant.*
- 11: SP-rodent.*
- 12: SP-virus.*
- 13: SP-vertebrate.*
- 14: SP-unclassified.*
- 15: SP-rvirus.*
- 16: SP-bacteriap.*
- 17: SP-archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2743	100.0	529	Q9AXK9	Q9axk9 adonis pala
2	2728	99.5	529	Q9AXL0	Q9axl0 adonis pala
3	1966.5	71.7	524	Q9LDV8	Q9ldv8 arabidopsis
4	1934	70.5	516	Q9FV43	Q9fv43 tagetes ere
5	1934	70.5	533	Q9AXK7	Q9axk7 lactuca sat
6	1892	69.0	517	Q8VWR6	Q8vwr6 spinacia ol
7	1786	65.1	437	Q8S3J4	Q8s3j4 citrus para
8	1518.5	55.4	382	Q9AXK8	Q9axk8 solanum tub
9	1113	40.6	262	Q8WLB9	Q8wlb9 citrus sine
10	783.5	28.6	504	Q9M546	Q9m546 citrus sine
11	778.5	28.4	524	Q9XGX3	Q9xgx3 citrus para
12	775.5	28.3	502	Q9AXL1	Q9axl1 adonis pala
13	755	27.5	511	Q9FV42	Q9fv42 tagetes ere
14	736	26.8	165	Q9SDF5	Q9sdf5 tagetes ere
15	729	26.6	495	Q8S3C3	Q8s3c3 sandersonia
16	719.5	26.2	498	Q9FV32	Q9fv32 lycopersico

17	714.5	26.0	498	10	Q9M424	Q9m424 solanum tub
18	710.5	25.9	498	10	Q9LWA6	Q9lwa6 lycopersico
19	499	18.2	201	10	Q9M5J5	Q9m5j5 daucus caro
20	288	10.5	165	10	Q9SDS5	Q9sds5 daucus caro
21	246	9.0	165	10	Q9SDS6	Q9sds6 daucus caro
22	173	6.3	394	2	Q9KIX3	Q9kix3 bradyrhizob
23	151.5	5.5	385	2	Q93CI8	Q93ci8 xanthobacte
24	151.5	5.5	394	16	Q9K157	Q9k157 neisseria m
25	146.5	5.3	394	16	Q9JSS1	Q9jss1 neisseria m
26	145.5	5.3	386	2	Q8VUJ6	Q8vu16 pentose agg
27	134.5	4.9	408	16	O06427	O06427 mycobacteri
28	134	4.9	393	17	Q8U4J0	Q8u4j0 pyrococcus
29	130	4.7	389	16	Q9CK99	Q9ck99 pasteurella
30	123	4.5	393	17	Q9V2B0	Q9v2b0 pyrococcus
31	123	4.5	530	10	Q9T064	Q9t064 arabidopsis
32	116.5	4.2	408	16	Q9CBA6	Q9cba6 mycobacteri
33	116.5	4.2	457	2	Q54453	Q54453 streptococc
34	115.5	4.2	299	17	Q8TV39	Q8tv39 methanopyru
35	115.5	4.2	379	16	Q8YOA5	Q8yoa5 raistonia s
36	115	4.2	704	16	Q8YRC4	Q8yrc4 anabaena sp
37	113	4.1	382	2	P94791	P94791 flavobacter
38	113	4.1	476	17	O59547	O59547 pyrococcus
39	112.5	4.1	405	16	Q9RI55	Q9ri55 streptomyce
40	112.5	4.1	1613	17	Q58907	Q58907 methanococc
41	112	4.1	413	2	Q9EXK5	Q9exk5 streptomyce
42	111.5	4.1	452	16	Q8YC38	Q8yc38 bruceella me
43	111.5	4.1	456	17	Q9HS24	Q9hs24 halobacteri
44	111.5	4.1	550	16	Q9KPA4	Q9kpa4 vibrio chol
45	110.5	4.0	381	17	O27753	O27753 methanobact

ALIGNMENTS

RESULT 1

ID	Q9AXK9	PRELIMINARY;	PRT;	529 AA.
AC	Q9AXK9;			
DT	01-JUN-2001 (TEMBLrel. 17, Created)			
DT	01-JUN-2001 (TEMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TEMBLrel. 19, Last annotation update)			
DE	Lycopene epsilon-cyclase.			
OS	Adonis palaeatina.			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;			
OC	Ranunculaceae; Adonis.			
OX	NCBI_TaxID=151078;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=IMMATURE AND DEVELOPING FLOWER BUDS;			
RX	MEDLINE=21126984; PubMed=11226339;			
RA	Cunningham F.X. Jr., Gantt E.;			
RT	"One ring or two? Determination of ring number in carotenoids by			
RT	lycopene varepsilon epsilon-cyclases";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:2905-2910(2001).			
CC	-!- COFACTOR: FAD (BY SIMILARITY).			
DR	EMBL; AF321536; AAK07432.1;			
DR	InterPro: IPR001327; FAD_pyr_redox.			
DR	PRINTS; PR00368; FADPNR.			
KW	FAD; Flavoprotein; Oxidoreductase.			
SQ	SEQUENCE 529 AA; 59170 MW; 784E4468F3D172B CRC64;			

Query Match 100.0%; Score 2743; DB 10; Length 529;
Best Local Similarity 100.0%; Pred No. 1.3e-221;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MELLGVRLNLISSCPVWTFGTFRNLSSKSLAYNIHRYGSSCRVDFQVRADGGSGSSVAYK 60	
Db	1	MELLGVRLNLISSCPVWTFGTFRNLSSKSLAYNIHRYGSSCRVDFQVRADGGSGSSVAYK 60	
QY	61	EGFVDEEDFTKAGSELFLVQMOTKSMERQAKLADKLPPIPFGESVMDLVVWICGCPAGL 120	
Db	61	EGFVDEEDFTKAGSELFLVQMOTKSMERQAKLADKLPPIPFGESVMDLVVWICGCPAGL 120	


```
QY 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIHAWKDTIVYLDNDAPVL 180
DB 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
DB 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
QY 241 VASGAASGKLLLEYEVGGPRVCVQTAYGVEVEENNPDPNLMVMDYRDYMQOQLQCSEE 300
DB 241 VASGAASGKLLLEYEVGGPRVCVQTAYGVEVEENNPDPNLMVMDYRDYMQOQLQCSEE 300
QY 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPDCLKRLKMSRLKTLGIQVTKVYEESYI 360
DB 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPDCLKRLKMSRLKTLGIQVTKVYEESYI 360
QY 361 PVGSLPNTQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
DB 361 PVGSLPNTQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
QY 421 QSSAVNTSMQAWSSLWPKEKRRQRAFFLGLLELIVQLDIEATRTFFRFLPTMMWGF 480
DB 421 QSSAVNTSMQAWSSLWPKEKRRQRAFFLGLLELIVQLDIEATRTFFRFLPTMMWGF 480
QY 481 LGSLSLSSFDLVLFMYMFLAPNSMRMSLVRLHLLSDPSGAVMVYAYLER 529
DB 481 LGSLSLSSFDLVLFMYMFLAPNSMRMSLVRLHLLSDPSGAVMVYAYLER 529

RESULT 2
Q9AXL0 PRELIMINARY; PRT: 529 AA.
ID Q9AXL0
AC Q9AXL0
DT 01-JUN-2001 (T-REMBLrel. 17, Created)
DT 01-JUN-2001 (T-REMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (T-REMBLrel. 19, Last annotation update)
DE Lycopene epsilon-cyclase.
OS Adonis palaeatina.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC Ranunculaceae; Adonis.
OX NCBI_TaxID=151078;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=IMMATURE AND DEVELOPING FLOWER BUDS;
RX MEDLINE=21126984; PubMed=11226339;
RA Cunningham F.X. Jr., Gantt E.;
RT "One ring or two? Determination of ring number in carotenoids by
RL lycopene varepsilon epsilon-cyclases.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:2905-2910(2001).
CC -1- COFACTOR: FAD (BY SIMILARITY).
DR EMBL: AF321535; AAK07431.1;
DR InterPro: IPR001327; FAD_pyr_redox.
DR PRINTS: PR00368; FADPNR.
KW FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 529 AA; 59155 MW; F44597CB0A239F9B CRC64;

Query Match 99.5%; Score 2728; DB 10; Length 529;
Best Local Similarity 99.1%; Pred. No. 2.3e-220;
Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVNRLISSCPVWTFGRNLSSSLKLYNHRIGSSCRVDFQVRADGGSGSRSSVAYK 60
DB 1 MELLGVNRLISSCPVWTFGRNLSSSLKLYNHRIGSSCRVDFQVRADGGSGSRSSVAYK 60
QY 61 EGFVDEEDFTKAGSELFLVQMOTKMEKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
DB 61 EGFVDEEDFTKAGSELFLVQMOTKMEKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
QY 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIHAWKDTIVYLDNDAPVL 180
DB 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIHAWKDTIVYLDNDAPVL 180
```

```
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
DB 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
QY 241 VASGAASGKLLLEYEVGGPRVCVQTAYGVEVEENNPDPNLMVMDYRDYMQOQLQCSEE 300
DB 241 VASGAASGKLLLEYEVGGPRVCVQTAYGVEVEENNPDPNLMVMDYRDYMQOQLQCSEE 300
QY 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPDCLKRLKMSRLKTLGIQVTKVYEESYI 360
DB 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPDCLKRLKMSRLKTLGIQVTKVYEESYI 360
QY 361 PVGSLPNTQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
DB 361 PVGSLPNTQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
QY 421 QSSAVNTSMQAWSSLWPKEKRRQRAFFLGLLELIVQLDIEATRTFFRFLPTMMWGF 480
DB 421 QSSAVNTSMQAWSSLWPKEKRRQRAFFLGLLELIVQLDIEATRTFFRFLPTMMWGF 480
QY 481 LGSLSLSSFDLVLFMYMFLAPNSMRMSLVRLHLLSDPSGAVMVYAYLER 529
DB 481 LGSLSLSSFDLVLFMYMFLAPNSMRMSLVRLHLLSDPSGAVMVYAYLER 529

RESULT 3
Q9LDV8 PRELIMINARY; PRT: 524 AA.
ID Q9LDV8
AC Q9LDV8
DT 01-OCT-2000 (T-REMBLrel. 15, Created)
DT 01-OCT-2000 (T-REMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (T-REMBLrel. 21, Last annotation update)
DE Lycopene epsilon cyclase (Putative lycopene epsilon cyclase).
GN AT5G57030.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
RN [2]
SEQUENCE FROM N.A.
RA Giuliano G., Rosati C., Santangelo G.;
RT "Gene structure and regulation of the carotenoid biosynthesis pathway
RT in Arabidopsis thaliana.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene MHM17.16/AT5G57030 (GI:8777443).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
```



```

DR PRINTS; PR00368; FADPNR.
KW FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 533 AA; 59832 MW; 7A5806BE758AE7B2 CRC64;

Query Match 70.5%; Score 1934; DB 10; Length 533;
Best Local Similarity 68.3%; Pred. No. 1.1e-153;
Matches 375; Conservative 57; Mismatches 77; Indels 40; Gaps 5;

QY 1 MELLGVRLN-----ISSCPVMT-----FGTRNLSSSLAYNIHRYGSSCR 40
DB 1 MECFGARNMATMAVFCPRFTDCNIRKPSLLKQRFNLSASSSLRQIKCSAKSDRCV 60

QY 41 VDFOVADGGGSSRSVAYKEGF--VDEEDFIKAGSELFLVQMOQTKSMKQAKLADKL 98
DB 61 VD-----KOGISVADEEDYKAGGSELFLVQMOQTKSMBSQSKLSEKL 103

QY 99 PPIPFESVMDLVVIGGPGAGLSAAEAALGLKVLGIGLDPDPTNNYGVWEDEFKDLGL 158
DB 104 AQPIGNCILDLVIGGPGAGLAAEASAKLGLNVLGIGLDPDPTNNYGVWQDEFGLGL 163

QY 159 ERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITE 218
DB 164 EGCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLSSKVERITE 223

QY 219 AGDGHSLWCENEIIFPCRLATVAGSAAGKLLLEYEGGPRVCVQTAIGVEVEYENNPYD 278
DB 224 APNGYSLIECEGNITPCRLATVAGSAAGKFLLEYELGGPRVCVQTAIGVEVEYENNPYD 283

QY 279 PNLVMDYRDYMOOKLQCESEETFLYVMPMSPTLRFEEETCLASKDAMPFDLLKRL 338
DB 284 PDLVMDYRDFSKHPESELEAKYPTFLYVMPMSPTLRFEEETCLASKEAMPFNLLSKL 343

QY 339 MSRLKTLGIVQTKVYEEESYIPVGGSLPNTQKNLAFAGAAASVMVHPATGYSVVRSLSA 398
DB 344 MSRLKANGIRITRYEEESYIPVGGSLPNTQKNLAFAGAAASVMVHPATGYSVVRSLSA 403

QY 399 PKYASVTAKILKQDNSAYVYVGGSSAVNISQWSSLPWPKRQRQRAFFLGLVOLD 458
DB 404 PNYAAVTAKILRODOSKEMIS-LGKYTNISKQAWETLWPLERKRQRQRAFFLGLSHVIMD 462

QY 459 IEARTFRFFRFLPTMMWGLGSSLSFDLVLFVYMFVLPAPNSMRMSLVRLHSDPS 518
DB 463 LEGTRFRFFRFLPKMMWGLGSSLSFDLVLFVYMFVLPAPNSMRMSLVRLHSDPT 522

QY 519 GAVNVRAYL 527
DB 523 GATWVKAYL 531

RESULT 6
Q8VWR6 PRELIMINARY; PRT; 517 AA.
AC Q8VWR6
DB Q8VWR6
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Lycopen epsilon-cyclase.
GN LEC.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MELODY; TISSUE=LEAF;
RA Desouza M.L., Kollmann S.R., Schroeder W.A.;
RT "Production of Lutein in Microorganisms.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF463497; AAL69394.1;
SQ SEQUENCE 517 AA; 57911 MW; 0B0EABB09376F8A4 CRC64;

Query Match 69.0%; Score 1892; DB 10; Length 517;

Best Local Similarity 71.4%; Pred. No. 3.6e-150;
Matches 355; Conservative 61; Mismatches 75; Indels 6; Gaps 1;

QY 37 SSCRVDFOVADGGGSSRS-----VAYKEGFVDEEDFIKAGSELFLVQMOQTKSMK 90
DB 19 SNVVESSRRSGSVLSANSDDSCVIAPEFANEEDFIKAGSELFLVQMOQKAMDC 78

QY 91 OAKLADKLPIPFESVMDLVVIGGPGAGLSAAEAALGLKVLGIGLDPDPTNNYGVWE 150
DB 79 YKISDKLRLQISDANELLDMVIGGPGAGLAAEASAKLGLKVLGIGLDPDPTNNYGVWE 138

QY 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 210
DB 139 DEFALGLGGCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 198

QY 211 SKVERITEAGDGHSLWCENEIIFPCRLATVAGSAAGKLLLEYEGGPRVCVQTAIGVEV 270
DB 199 AKVENIMEGPDGRLVACERGVTPCRLVTVAGSAAGKLLLEYEGGPRVCVQTAIGVEV 258

QY 271 EVENNPYDNLVMDYRDYMOOKLQCESEETFLYVMPMSPTLRFEEETCLASKDAMP 330
DB 259 EVENSYPDNPVMDYRDYTKLSVQSLKAYPTFLYAMPISPTRIFFEEETCLASVDAMP 318

QY 331 FDLKRLKMSRLKTLGIVQTKVYEEESYIPVGGSLPNTQKNLAFAGAAASVMVHPATGY 390
DB 319 FDLKRLKMLTRLOTWGRITRYEEESYIPVGGSLPNTQKNLAFAGAAASVMVHPATGY 378

QY 391 VVRSLSAPKVASVTAKILKQDNSAYVYVGGSSAVNISQWSSLPWPKRQRQRAFFLFG 450
DB 379 VVRSLSAPKVASVTAKILKQDNSAYVYVGGSSAVNISQWSSLPWPKRQRQRAFFLFG 438

QY 451 LELIVQDLIEATRTFFRFLPTMMWGLGSSLSFDLVLFVYMFVLPAPNSMRMSLV 510
DB 439 LSLIVQDLIEGIRTFRFFRFLPKMMWGLGSSLSADLILFAYMFVLPAPNSMRMSLV 498

QY 511 RHLSDPSGAVNVRAYL 527
DB 499 RHLSDPTGATMIRTY 515

RESULT 7
Q8S3J4 PRELIMINARY; PRT; 437 AA.
AC Q8S3J4
DB Q8S3J4
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Lycopen epsilon-cyclase.
OS Citrus paradisi (Grapefruit).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=37656;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DUNCAN;
RA Costa M.C., Otoni W.C., Moore G.A.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486650; AAL92114.1;
SQ SEQUENCE 437 AA; 49012 MW; 5BECC80EF76EE4 CRC64;

Query Match 65.1%; Score 1786; DB 10; Length 437;
Best Local Similarity 77.4%; Pred. No. 2.2e-141;
Matches 333; Conservative 44; Mismatches 53; Indels 0; Gaps 0;

QY 98 LPPIPFESVMDLVVIGGPGAGLSAAEAALGLKVLGIGLDPDPTNNYGVWEDEFKDLG 157
DB 6 LPPISIGNGILDLVIGGPGAGLAAEASAKLGLNVLGIGLDPDPTNNYGVWEDEFKDLG 65

QY 158 LERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERIT 217
DB 66 LEGCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLSSKVESIT 125

```

QY 218 EAGDGHSLVVCNEIFIPCLATVAGSAGSKLLEYEGGPRVCVQATAYGVEVEENNPY 277
 Db 126 ESTSGHRLVACSHDMLPCLATVAGSAGSKLLEYEGGPRVCVQATAYGVEVEENNPY 185
 QY 278 DPNLAVFMDYRDYMOOKLQCSSEETLYVMPMSPTFLFFETCLASKDAMPDILLKK 337
 Db 186 DPNLAVFMDYRDYMOOKLQCSSEETLYVMPMSPTFLFFETCLASKDAMPDILLKK 245
 QY 338 LMSRLKTLGIQVTKYEEWSPYIPVGGSLPNTFQKNLAFGAASVMVHPATGYSVVRSLSE 397
 Db 246 LMSRLKTLGIQVTKYEEWSPYIPVGGSLPNTFQKNLAFGAASVMVHPATGYSVVRSLSE 305
 QY 398 APKASVIAKILKQNSAVVSGQSAVNISMQAQSSWLPKPKRQRAFFLGLLEIVOL 457
 Db 306 APKASVIAKILKQNSAVVSGQSAVNISMQAQSSWLPKPKRQRAFFLGLLEIVOL 365
 QY 458 DEARTFTFTRPFLPNTFQKNLAFGAASVMVHPATGYSVVRSLSE 517
 Db 366 DEARTFTFTRPFLPNTFQKNLAFGAASVMVHPATGYSVVRSLSE 425
 QY 518 SGAVMVRAYL 527
 Db 426 TCGATWRTYL 435

RESULT 8
 Q9AXK8 PRELIMINARY; PRT; 382 AA.
 AC Q9AXK8;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Lycopene epsilon-cyclase (Fragment).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21126984; PubMed=11226339;
 RA Cunningham F.X. Jr., Gantt E.;
 RT "One ring or two? Determination of ring number in carotenoids by
 lycopene varepsilon epsilon-cyclases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:2905-2910(2001).
 DR EMBL; AF321537; AAK07433.1; -
 FT NON_TER 1
 SQ SEQUENCE 382 AA; 43249 MW; E9490FB56DC7D710 CRC64;

Query Match 55.4%; Score 1518.5; DB 10; Length 382;
 Best Local Similarity 74.5%; Pred. No. 5.1e-119;
 Matches 281; Conservative 48; Mismatches 47; Indels 1; Gaps 1;
 QY 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGSYLD 210
 Db 5 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGSYLD 64
 QY 211 SKVERITEAGDGHSLVVCNEIFIPCLATVAGSAGSKLLEYEGGPRVCVQATAYGVEV 270
 Db 65 SKVDRIVEATNGHSLVCEGDVIPCFTVTVASGASGKFLQYELGGPRVSVQATAYGVEV 124
 QY 271 EVNNPYPNLMVMDYRDYMOOKLQCSSEETLYVMPMSPTFLFFETCLASKDAMP 330
 Db 125 EVNNPYPNLMVMDYRDYMOOKLQCSSEETLYVMPMSPTFLFFETCLASKDAMP 184
 QY 331 FOLLKRLKMSRLKTLGIQVTKYEEWSPYIPVGGSLPNTFQKNLAFGAASVMVHPATGYS 390
 Db 185 FOLLKRLKMSRLKTLGIQVTKYEEWSPYIPVGGSLPNTFQKNLAFGAASVMVHPATGYS 244
 QY 391 VYRSLEAPKASVIAKILKQNSAVVSGQSAVNISMQAQSSWLPKPKRQRAFFLGL 450
 Db 245 VYRSLEAPKASVIAKILKQNSAVVSGQSAVNISMQAQSSWLPKPKRQRAFFLGL 303

QY 451 LELIOLDIEARTFTFTRPFLPNTFQKNLAFGAASVMVHPATGYSVVRSLSE 510
 Db 304 LELIOLDIEARTFTFTRPFLPNTFQKNLAFGAASVMVHPATGYSVVRSLSE 363
 QY 511 RHLLSDPSCAVMVRAYL 527
 Db 364 RHLLSDPSCAVMVRAYL 380

RESULT 9
 Q9W1B9 PRELIMINARY; PRT; 262 AA.
 AC Q9W1B9;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Lycopene epsilon-cyclase (Fragment).
 OS Citrus sinensis (Sweet orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=2711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu C., Qin Q., Zhang S.;
 RT "Molecular cloning of lycopene epsilon-cyclase gene from orange
 (Citrus sinensis).";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF450280; AAL47019.1; -
 FT NON_TER 1
 SQ SEQUENCE 262 AA; 29104 MW; E19FB6A5D325D8DD CRC64;

Query Match 40.6%; Score 1113; DB 10; Length 262;
 Best Local Similarity 79.0%; Pred. No. 3.2e-85;
 Matches 207; Conservative 27; Mismatches 28; Indels 0; Gaps 0;
 QY 120 LSLAAEAAKLGKLVGLIGDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIVYLDNDAPV 179
 Db 1 LSLAAEAAKLGKLVGLIGDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIVYLDNDAPV 60
 QY 180 LIGRAYGRVSRHLLHEELLKRCVSGSYLDKVERITEAGDGHSLVVCNEIFIPCLRLA 239
 Db 61 LIGRAYGRVSRHLLHEELLKRCVSGSYLDKVERITEAGDGHSLVVCNEIFIPCLRLA 120
 QY 240 TVASGAASGKLLLEYEGGPRVCVQATAYGVEVEENNPYPNLMVMDYRDYMOOKLQCS 299
 Db 121 TVASGAASGKLLLEYEGGPRVCVQATAYGVEVEENNPYPNLMVMDYRDYMOOKLQCS 180
 QY 300 EYPTFLYVMPMSPTFLFFETCLASKDAMPFOLLKRLKMSRLKTLGIQVTKYEEWSPY 359
 Db 181 SDNPTFLYVMPMSPTFLFFETCLASKDAMPFOLLKRLKMSRLKTLGIQVTKYEEWSPY 240
 QY 360 IPVGGSLPNTFQKNLAFGAAS 381
 Db 241 IPVGGSLPNTFQKNLAFGAAS 262

RESULT 10
 Q9M546 PRELIMINARY; PRT; 504 AA.
 AC Q9M546;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Lycopene beta-cyclase.
 OS Citrus sinensis (Sweet orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=2711;
 RN [1]

```

RP SEQUENCE FROM N.A.
RA Xu C.J., Zhang S.L.;
RT "Molecular cloning of lycopene beta-cyclase gene from orange (Citrus
RL sinensis).";
DR EMBL; AF240787; AAF44700.2; -
DR InterPro; IPR00205; NAD_binding.
SQ SEQUENCE 504 AA; 56460 MW; 7C12B2BFB646481C CRC64;

Query Match      28.6%; Score 783.5; DB 10; Length 504;
Best Local Similarity 38.3%; Pred. No. 4.1e-57;
Matches 172; Conservative 94; Mismatches 160; Indels 23; Gaps 10;

QY 69 FIKAGGSELL-FVQMOTKSMKQAKLADKLPPFPFGESVMDLVVIGCGPAGLSAAEAA 127
DB 53 FIKASSSALLELVPEETKKELEFELPMYD-----PSKGLVVDLAVVGGGPAGLVAQVVS 107
QY 128 KLGLKVLGI--GPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAY 185
DB 108 EAGLSVCSIDSPKLIWPNNTYGVWVDEFEAMDLLDCLDTTWSGAVVHDDDTKKDLDRPY 167
QY 186 GRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCRLATVASGA 245
DB 168 GRVNRKLLKSMLOKQICITNGVKFHOAKVIKVIHE-ESKSLICNDGVTIQAQVLDATG- 225
QY 246 ASGKLLEYEVG-GRVVCVQTAYGVEVENNPDNLMVFMDYRD---YMOOKLQCSSEE 301
DB 226 FSRCLVQYDKPYNPG--YQVAYGILAEVEHPFDLDMKVFMDWRDLSHNNSELKEANSK 283
QY 302 YPTFLYVMPSPTRLPFEETCLASKDAMPFDLLKRLKMSRLTKLGIOVTKVYEEWSYIP 361
DB 284 IPTFLYAMPFSNRIFFLEETSLVARPGVPMKDIQERMVARLKHGKIKVRSIEDEHCVIP 343
QY 362 VGGSLPNTQKNLAFGAASVMHPATGYSVVRSLSSEAPKYASVIAKILKQNSAYVVSQ 421
DB 344 MGGPLVPLPQRVGIGGTAGMVHPSTGVMVARTLAAPIVANAIVRSLSRSRDS---ISGH 400
QY 422 SSAYNISQAWSSLLWPKERKQRAFFLGLLEIVOLDIEATRTFTFRFLPTMMWGF 481
DB 401 ----KLSAEVWKDLWPIERRRQREFFCFGMDILLKLDLPATRRFFDAFDELPYRWGHFL 456
QY 482 GSSLSFDLVLFMSYMFVLAIPNSMRMSLV 510
DB 457 SSRFLPELVGLSLFASHASNTSRLEIM 485

RESULT 11
QYXGX3
ID QYXGX3 PRELIMINARY; PRT; 524 AA.
AC QYXGX3;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Lycopene cyclase.
OS Citrus paradisi (Grapefruit).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=37656;
RN [1]
RP SEQUENCE FROM N.A.
RA Costa M.C., Moreira C.D., Melton J.R., Otoni W.C., Moore G.A.;
RT "Developmental expression of carotenoid genes in Citrus.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152246; AAD38049.2; -
DR InterPro; IPR00205; NAD_binding.
SQ SEQUENCE 524 AA; 58523CF5B3B9208C CRC64;

Query Match      28.4%; Score 778.5; DB 10; Length 524;
Best Local Similarity 38.3%; Pred. No. 1.1e-56;
Matches 172; Conservative 93; Mismatches 161; Indels 23; Gaps 11;

QY 69 FIKAGGSELL-FVQMOTKSMKQAKLADKLPPFPFGESVMDLVVIGCGPAGLSAAEAA 127

```

```

DB 73 FIKASSSALLELVPEETKKELEFELPMYD-----PSKGLVVDLAVVGGGPAGLVAQVVS 127
QY 128 KLGLKVLGI--GPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAY 185
DB 128 GAGLSVCSIDSPKLIWPNNTYGVWVDEFEAMDLLDCLDTTWSGAVVHDDDTKKDLNRPY 187
QY 186 GRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCRLATVASGA 245
DB 188 GRVNRKLLKSMLOKQICITNGVKFHOAKVIKVIHE-ESKSLICNDGVTIQAQVLDATG- 245
QY 246 ASGKLLEYEVG-GRVVCVQTAYGVEVENNPDNLMVFMDYRD-YMOOKLQCSSEE--EE 301
DB 246 FSRCLVQYDKPYNPG--YQVAYGILAEVEHPFDLDMKVFMDWRDLSHNNSELKEANSK 303
QY 302 YPTFLYVMPSPTRLPFEETCLASKDAMPFDLLKRLKMSRLTKLGIOVTKVYEEWSYIP 361
DB 304 IPTFLYAMPFSNRIFFLEETSLVARPGVPMKDIQERMVARLKHGKIKVRSIEDEHCVIP 363
QY 362 VGGSLPNTQKNLAFGAASVMHPATGYSVVRSLSSEAPKYASVIAKILKQNSAYVVSQ 421
DB 364 MGGPLVPLPQRVGIGGTAGMVHPSTGVMVARTLAAPIVANAIVRSLSRSRDS---ISGH 420
QY 422 SSAYNISQAWSSLLWPKERKQRAFFLGLLEIVOLDIEATRTFTFRFLPTMMWGF 481
DB 421 ----KLSAEVWKDLWPIERRRQREFFCFGMDILLKLDLPATRRFFDAFDELPYRWGHFL 476
QY 482 GSSLSFDLVLFMSYMFVLAIPNSMRMSLV 510
DB 477 SSRFLPELVGLSLFASHASNTSRLEIM 505

RESULT 12
QYXAXL1
ID QYXAXL1 PRELIMINARY; PRT; 502 AA.
AC QYXAXL1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Lycopene beta-cyclase.
OS Adonis palaeatina.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC Ranunculaceae; Adonis.
OX NCBI_TaxID=151078;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21126984; PubMed=11226339;
RA Cunningham F.X. Jr., Gantt E.;
RT "One ring or two? Determination of ring number in carotenoids by
RL lycopene varepsilon epsilon-cyclases.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:2905-2910(2001).
DR EMBL; AF321534; AAK07430.1; -
DR InterPro; IPR00205; NAD_binding.
SQ SEQUENCE 502 AA; 56575 MW; B973FD43794D3F1B CRC64;

Query Match      28.3%; Score 775.5; DB 10; Length 502;
Best Local Similarity 36.3%; Pred. No. 1.9e-56;
Matches 181; Conservative 99; Mismatches 166; Indels 53; Gaps 14;

QY 23 LSSSKLAYNIHRYGS-----SCRVDQVRADGGSGSRSSVAYKEGFVDEEDFIKAGSELL 78
DB 27 VSTSKLQNVFRIASRNTHPCR-----NGT-----VKARGSALL 60
QY 79 FVQMOTKSMKQAKLADKLPPFPFGESVMDLVVIGCGPAGLSAAEAAKILKVLGI- 136
DB 61 ----ELVPETKKELEFDPAYDPSRGIVVDLAVVGGGPAGLTAQVSEAGLVCSID 115
QY 137 -GPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHE 195
DB 116 FSPKLIWPNNTYGVWVDEFEAMDLLDCLDTTWSGAVVYTDNNSKKYLDRIYGRVNRKQLKS 175
QY 196 ELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCRLATVASGAASGKLLEYEV 255

```

```

Db 176 KMLQKVTNGVHFHQAQKIVKVIHE-ESKSLICNDGITTNAIVLVDATG-FSRCLVQYDK 233
QY 256 G-GPRVCVCTAGYGEVEENNNPDNLMVFMVDRD-YMOOKLQCSSE--EYTFELVVMW 311
Db 234 PYNPG--YQVAGINAEVEHEFDLKLMLFMDWRSHLNEKLELDKNRKIPITFLYAMPF 291
QY 312 STRLFFETCTCLASKADAMPFDLLKRLMSRLKTLGIQVTKVYEEWYSYIPVGGSLPNTQ 371
Db 292 SSTKIFLETSVLARPGFLAFQIERMAYRLKHLGIKVKISIEDERCVCIPMGGLPVLQ 351
QY 372 KNLAGAASVHPATGYSVRSLSSEAPKYASVIAKILKQDMSAYVVSQSSAVNISMQA 431
Db 352 RVVGIGTAGMHPSTGYMARTLAAPVVAKSIVQYLGSDRS---LSGN----ELSAEV 404
QY 432 WSLMPKRRQRAFFLFLGLELIVQIDIEATRTFFTRFLPTMMWGLSSLSFSLV 491
Db 405 WKDLPIERRRQREFFCFGMDILLKLDLGTRRFFDAFFDLPHYWHGFLSKRLFLPELL 464
QY 492 LFSMTMFLVAPNSMRSLV 510
Db 465 FFGLSLFSNASRIEIM 483

```

```

RESULT 13
Q9FV42
ID Q9FV42 PRELIMINARY; PRT; 511 AA.
AC Q9FV42;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Beta cyclase.
OS Tagetes erecta (African marigold).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroidae;
OC Heleniaceae; Tagetes.
OX NCBI_TaxID=13708;
RN [1]
RP SEQUENCE FROM N.A.
RA Moehs C.P., Tian L., DellaPenna D.;
RT "Analysis of carotenoid biosynthetic gene expression during marigold
RT petal development.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -I- COFACTOR: FAD (BY SIMILARITY).
DR EMBL; AF251017; AAG10429.1;
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR001100; Pyr_redox.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNRDRTASEI.
KW FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 511 AA; 57831 MW; C72CEF6C3F2EF163 CRC64;

```

```

Query Match 27.5%; Score 755; DB 10; Length 511;
Best Local Similarity 36.5%; Pred. No. 1e-54;
Matches 181; Conservative 97; Mismatches 188; Indels 30; Gaps 15;

QY 25 SKLAYNIHRYG--SSCRVDFOVRADGGSGSSSVAYKEGFVDEEDFIKAGSELL-FVQ 81
Db 17 SNKFGNLLNQLNQSKSQFO---DFRGPKKK--QKLG---QKVCVKAASSALLELYP 69
QY 82 MOOTKSMKQAKLADLPPIPFGEVMDLVVIGCGPAGLSLAAEAAKLGKLVGLI--GPD 139
Db 70 EIKKENLDFDLPWD-----PSRNVVDLVVVGSGSLAVAQVSEAGLTVCSDIDPSK 124
QY 140 LPFTNYGWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELK 199
Db 125 LIWPNNGYGVWDEFAEMDLLDCLDTTWSAVYIDEKSTKSLNRPYARNRQLKTKMLQ 184
QY 200 RCVESGVSLDSKVRITEAGDGHSLVCENEIFIPCLRLATVASGAASKLLEYEGV-GP 258
Db 195 KCIANGVHFHQAQKIVKVIHE-ELKSLICNDGVTIQTATLVLDATG-FSRSLVQYDKPYNP 242

```

```

QY 259 RVCVQTAGYGEVEENNNPDNLMVFMVDRD-YMOOKLQ--CSEEBVPTFLYVMPKSPTR 315
Db 243 G--YQVAYGILAEVEHEFPDVKMLFMDWRDSDLQNLKARNSRIPITFLYAMPFSSTR 300
QY 316 LFFETCTCLASKADAMPFDLLKRLMSRLKTLGIQVTKVYEEWYSYIPVGGSLPNTQKLA 375
Db 301 IFLEETSILVARPGKLMEDIQERMAYRLKHLGIKVKISIEDERCVCIPMGGLPVLQ 360
QY 376 FGAASVHPATGYSVRSLSSEAPKYASVIAKILKQDMSAYVVSQSSAVNISMQA 434
Db 361 IGTAGMHPSTGYMARTLAAPVVAKSIVQYLGSDRS---LSGN----ELSAEV 416
QY 435 LMPKRRQRAFFLFLGLELIVQIDIEATRTFFTRFLPTMMWGLSSLSFSLV 494
Db 417 LMPIERRRQREFFCFGMDILLKLDLGTRRFFDAFFDLPHYWHGFLSKRLFLPELL 476
QY 495 MYMFLVAPNSMRSLV 510
Db 477 LSLFGHASNTCRVEIM 492

```

```

RESULT 14
Q9SDP5
ID Q9SDP5 PRELIMINARY; PRT; 165 AA.
AC Q9SDP5;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE Lycopene epsilon-cyclase (fragment).
GN LYC-E.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen D., Chen H., Wang Y.;
RT "Daucus carota lycopene epsilon-cyclase.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212130; AAF23013.1;
FT NON_TER 1
FT NON_TER 165
SQ SEQUENCE 165 AA; 18455 MW; F61CDAB0B4B53861 CRC64;

```

```

Query Match 26.8%; Score 736; DB 10; Length 165;
Best Local Similarity 83.6%; Pred. No. 7.1e-54;
Matches 138; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 146 YGVWDEDFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELKRCVESG 205
Db 1 YGVWDEFDLGLGEGIEHWRDTIVYLDGDPIMIGRAYGRVSRHLLHEELKRCVESG 60
QY 206 VSYLDSKVRITEAGDGHSLVCENEIFIPCLRLATVASGAASKLLEYEGVGRVCVQTA 265
Db 61 VSYLSSKVEKITEAGDGHSLVCENINIVIPCLRLATVASGAASKLLEYEGVGRVSVQTA 120
QY 266 YGVEVEENNNPDNLMVFMVDRDYMOKLQCSSEYPTFLYVMP 310
Db 121 YGVEVEENNNPDNLMVFMVDRDYMOKLQCSSEYPTFLYVMP 165

```

```

RESULT 15
Q8S3C3
ID Q8S3C3 PRELIMINARY; PRT; 495 AA.
AC Q8S3C3;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Beta-lycopene cyclase.
OS Sanderionia aurantiaca (Christmas-bells) (Chinese-lantern lily).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:17:23 ; Search time 2143 Seconds
(without alignments)
3997.859 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELGVRNLSSCPVTFGT.....VRHLLSPGAVMVRAYLER 529

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPPO_spool/US09701395/runat_15052003_165751_3370/app_query.fasta_1.711
-DB=EST -QFMT=fastap -SURFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09701395 -CGN_1_1306.runat_15052003_165751_3370 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6 -FCAPEXT=7
-YGAPOP=10 -YGAPEXT=10 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_rrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1048.5	38.2	755	14	BQ856148 QGB28L15
2	899	32.8	664	14	BQ623342 USDA-PP-0
3	868.5	31.7	658	12	BG592655 EST491333
4	860	31.4	661	14	BQ994803 QGF8D16.Y
5	843.5	30.8	719	10	AV931085 AV931085
6	793.5	28.9	665	10	AV926200 AV926200
7	728.5	26.6	610	14	BQ116144 EST601720
8	701	25.6	541	12	BG241188 OVI_39.CO
9	695	25.3	548	10	AW906884 EST343007
10	619.5	22.6	571	10	BE231430 SsS0941.S
11	602.5	22.0	614	10	AV921185 AV921185
12	600	21.9	557	10	AV521661 AV521661
13	599	21.8	465	10	AW982124 SsS0040.S
14	598.5	21.8	668	10	AV822597 AV822597
15	581.5	21.2	617	13	BJ482945 BJ482945
16	579.5	21.1	512	10	AV916064 AV916064
17	575	21.0	652	13	BI096991 SCUM33-DS
18	545	19.9	871	10	BE456098 HVSMEg001
19	516	18.8	655	12	BG240869 OVI_39.CO
20	506.5	18.5	841	17	BH446159 BOGJF81TR
21	500.5	18.2	442	10	AV917261 AV917261
22	498	18.2	673	13	BM449133 DSA032D12
23	460	16.8	341	10	BE318207 NF036F05L
24	459.5	16.8	622	12	BF636131 NF079H03D
25	451	16.4	325	12	BF067883 ST84f04.Y
26	425.5	15.7	420	12	BF145063 SsS0752.S
27	420	15.3	599	13	BF1992439 1020062B0
28	406.5	14.8	430	10	AW599513 ga87e04.Y
29	394	14.4	547	12	BG126518 BG126518
30	390	14.2	692	12	BG127631 EST473193
31	384	14.0	590	9	AI485631 EST243952
32	375	13.7	443	14	BQ872792 QG116K13
33	374.5	13.7	743	17	BH598608 BOGVE83TF
34	368	13.4	381	14	BU012051 QGJ1701.0
35	354.5	12.9	573	12	BF423984 sso2q08.Y
36	335	12.2	649	13	BI921781 EST541684
37	334	12.2	854	14	BQ509489 EST616904
38	333.5	12.2	552	12	BG524479 42-95.Ste
39	330	12.0	487	17	BH233819 1006175H1
40	327	11.9	611	13	BI422789 EST533455
41	326.5	11.9	336	12	BF145076 SsS0766.S
42	322	11.7	404	10	AV803737 AV803737
43	319.5	11.6	507	14	BQ872037 QG113H23
44	319.5	11.6	526	14	BQ980868 QGEL12D21
45	318	11.6	601	12	BG646437 EST508056

ALIGNMENTS

RESULT 1
BQ856148
LOCUS
DEFINITION QGB28L15.Yg.ab1 QG-ABCDI lettuce salinas Lactuca sativa cDNA clone
ACCESSION QGB28L15, mRNA linear EST 14-AUG-2002
VERSION QGB56148, mRNA sequence.
KEYWORDS QGB56148.1 GI:22241613
SOURCE EST.
ORGANISM Lactuca sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
REFERENCE 1 (bases 1 to 755)

AUTHORS

Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevaller, P., Ziegler, J., Ellison, P., Kolkman, J., Siabang, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L., and Bradford, K.
 lettuce and Sunflower ESTs from the Compositae Genome Project
 Unpublished (2002)
<http://compenomics.ucdavis.edu/>

TITLE

lettuce and Sunflower ESTs from the Compositae Genome Project
 Unpublished (2002)
<http://compenomics.ucdavis.edu/>

JOURNAL

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmudson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.edu [michelmore@veggmail.ucdavis.edu]
 belongs to contig OG_CA_Contig1411, see <http://cgdb.ucdavis.edu/>
 for details.
 Plate: QGB28 row: L column: 15.

FEATURES

source

Location/Qualifiers
 1. .755
 /organism="Lactuca sativa"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="QGB28L15"
 /clone_lib="QG-ABCDI lettuce salinas"
 /lab_host="E.coli"

/note="Vector: pBRCDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>
 TAG_LIB-QG ABCDI lettuce salinas
 TAG_TISSUE=flowers post-fertilized
 TAG_SEQ=TCGCATCGG"

BASE COUNT 223 a 149 c 159 g 223 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 4,04e-124 Length: 755
 Score: 1048.50 Matches: 194
 Percent Similarity: 88.49% Conservative: 29
 Best Local Similarity: 76.98% Mismatches: 28
 Query Match: 38.22% Indels: 1
 DB: 14 Gaps: 1

US-09-701-395A-23 (1-529) x B0856148 (1-755)

QY 256 GlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluAsnAsn 275
 Db 1 GGGGGTCCCGTGTGTTGTCACCAACAGCTTATGTTAGAGGTGAGTTGAAACAAAC 60
 QY 276 ProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnLysLeu 295
 Db 61 CCTATGATCCAGATCAATTAAGGGTTCATGATATAGAGCTCTCAAAACATAAACCG 120
 QY 296 GlnCysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArg 315
 Db 121 GAATCTTTAGAACCAAAATATCCGACTTCTCTATGTCATGGCCATGTCACCAACAAA 180
 QY 316 LeuPhePheGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLys 335
 Db 181 ATATCTTTTGGAGAACTTCTAGCTTCAAGGAGCCATGCTTCAATCTCTAAAG 240
 QY 336 ArgLysLeuMetSerArgLeuLysThrLeuGlyTleGlnValThrLysValTyrGlu 355
 Db 241 TCCAAACTCATGTCACGATTAAAGGCAATGGGTATCCGAATAACAAAGACGTACGAAGAG 300
 QY 356 GluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluLysAsnLeuAla 375
 Db 301 GAATGGTTCGTATATCCCGTAGGTGGATCGTTACCTAATACAGAACAAAGATCTCGCA 360

QY 376 PheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeu 395
 Db 361 TTTGGTGTGCGACTAGTATGGTGCACCTGCGACAGGGTATTACAGTTGTCGATCTTG 420
 QY 396 SerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAla 415
 Db 421 TCAGAGCTCTAATATATGACGAGTCATTGCTAAGATTTTAAGACAGATCAATCTAAA 480
 QY 416 TyrValValSerGlyGlnSerAlaValAsnIleSerMetGlnAlaTtpSerSerLeu 435
 Db 481 GAAATGATTCTCT--CTTGGAAATACACTAACATTTCAAACAGCATGGGAACATG 537
 QY 436 TrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuVal 455
 Db 538 TGCCACTTCAAGAGGAAAGACAGCAGCCTCTCTTCTATTTCGGACTATCACACATCGT 597
 QY 456 GlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTyr 475
 Db 598 CTAATGGATCTAGAGGGAACACGTCACATTTTTCGCTACTTTCTTCGTTTCCCANATGG 657
 QY 476 MetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMet 495
 Db 658 ATGTGGTGGGATTTTGGGGTCTCTTCTTATCTTCAACGGATTTGATAATATTTCGGCTT 717
 QY 496 TyrMetPheValLeuAlaProAsnSerMetArgMet 507
 Db 718 TATATGTTTGTATAGCACCCTCACAGCTTGAGAAATG 753

RESULT 2

BQ623342

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1. .664

/organism="Citrus sinensis"

/cultivar="Ridge Pineapple"

/db_xref="taxon:2711"

/clone="USDA-FP_00433"

/clone_lib="Ridge pineapple sweet orange entire seedling"

/tissue_type="entire seedling"

/dev_stage="50 days after germination"

/lab_host="XL1-Blue"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; A high quality EST with at least 200 contiguous bases at Trace runner score of 20 or better"

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-701-395A-23 (1-529) x B0856148 (1-755)

QY 256 GlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluAsnAsn 275

Db 1 GGGGGTCCCGTGTGTTGTCACCAACAGCTTATGTTAGAGGTGAGTTGAAACAAAC 60

QY 276 ProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnLysLeu 295

Db 61 CCTATGATCCAGATCAATTAAGGGTTCATGATATAGAGCTCTCAAAACATAAACCG 120

QY 296 GlnCysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArg 315

Db 121 GAATCTTTAGAACCAAAATATCCGACTTCTCTATGTCATGGCCATGTCACCAACAAA 180

QY 316 LeuPhePheGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLys 335

Db 181 ATATCTTTTGGAGAACTTCTAGCTTCAAGGAGCCATGCTTCAATCTCTAAAG 240

QY 336 ArgLysLeuMetSerArgLeuLysThrLeuGlyTleGlnValThrLysValTyrGlu 355

Db 241 TCCAAACTCATGTCACGATTAAAGGCAATGGGTATCCGAATAACAAAGACGTACGAAGAG 300

QY 356 GluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluLysAsnLeuAla 375

Db 301 GAATGGTTCGTATATCCCGTAGGTGGATCGTTACCTAATACAGAACAAAGATCTCGCA 360

Pred. No.: 6.86e-105 Length: 664
 Score: 899.00 Matches: 172
 Percent Similarity: 87.84% Conservative: 23
 Best Local Similarity: 77.48% Mismatches: 24
 Query Match: 32.77% Indels: 3
 DB: 14 Gaps: 0

US-09-701-395A-23 (1-529) x BQ623342 (1-664)

Qy 75 SerGluLeuLeuPheValGluMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
 Db 3 TCTAGCTTTTTCGTAACATGAGCAAAACAGTCAATGGATAACAGTCTAAATTA 62
 Qy 95 AlaAspLysLeuProPheGlyGluSerValMetAspLeuValValIleGly 114
 Db 63 GCCGATAAGCTACGCCAATATCAATGGTATGTTTGGTATTTGGTGGTGGT 122
 Qy 115 CysGlyProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGly 134
 Db 123 TGTGGCCAGCTGCTGCTTTGGCTGCAATCAGCGAAGTTGGGATTAATTTGGA 182
 Qy 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTyrGluAspGluPheLys 154
 Db 183 CTTATTTGCCCGGATCTCCCTTTTCAACAAATATGTTGGGAGATGAATTTAGA 242
 Qy 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAsp 174
 Db 243 GATCTTGGACTTGAAGGTGTATCGAACATGCTCTGGAGACACAGTGTATATATTGAT 302
 Qy 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHis 194
 Db 303 GAAGTAACCCATCTGATTGGTGGCTTATGACAGTGTAGTCGACATTTGCTTCAT 362
 Qy 195 GluGluLeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGlu 214
 Db 363 GAAGATTTATTAAGAGGTGTCTGAGTCAGTGTTCATATCTTAGCTCAAAAGTGAA 422
 Qy 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle 234
 Db 423 AGCATTTACGGAATCTACCAAGTTCCTGCTGATGTTGTTGATGATGATGATGTC 482
 Qy 235 ProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuGluTyrGlu 254
 Db 483 CCTTCAGGCTGTCTACGTGCTCTCTGAGCAGCATCAGGAAGCTATTGGAATATGAG 542
 Qy 255 -ValGlyGlyProArgValCysValGlnThrAlaTyrGlyVal--GluValGluValGlu 273
 Db 543 GGGGGGGGTCCTCCAAAGTTCCTGCTCAACAGCTTATGTTGGAGGTTAAAGGGGAA 602
 Qy 274 AsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGln 293
 Db 603 AATAATCCATATGATCCAAAGCTTATGTTTCTATGGACTACAGAGACTGTACTAAGCAA 662

RESULT 3
 BG592655 658 bp mRNA linear EST 12-APR-2001
 DEFINITION BG592655 cSTS Solanum tuberosum cDNA clone cSTS1P24 5' sequence,
 mRNA sequence.

ACCESSION BG592655
 VERSION BG592655.1 GI:13610795
 KEYWORDS EST.
 SOURCE potato.

ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 658)
 van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
 Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
 Generations of ESTs from sprouting potato eyes
 Unpublished (2000)
 CONTACT: Cathy Ronning
 The Institute for Genomic Research

For clone info: please contact Research Genetics, Libraries
 Division tel 1-800-711-6195, email cdna@resgen.com
 Seq primer: M13P-R.

FEATURES

source
 1..658
 Location/Qualifiers
 /organism="Solanum tuberosum"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS1P24"
 /clone_lib="cSTS"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOL"

/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
 XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
 taken from tubers. The tubers were incubated at 26C in the
 dark for 2-3 weeks prior to sprouting. The eyes were
 frozen in liquid nitrogen immediately upon removal from
 tubers."

BASE COUNT 194 a 137 c 148 g 179 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5.92e-101 Length: 658
 Score: 868.50 Matches: 163
 Percent Similarity: 87.27% Conservative: 29
 Best Local Similarity: 74.09% Mismatches: 27
 Query Match: 31.66% Indels: 1
 DB: 12 Gaps: 1

US-09-701-395A-23 (1-529) x BG592655 (1-658)

Qy 226 ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla 245
 Db 2 GTAGATCGGAGGTGATGTGTGATTCCTCGCAGGTTGTCTGCTGTCATCGGGAGCA 61
 Qy 246 AlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGlnThrAla 265
 Db 62 GCCTCGGGAAATCTTCAGTATGAGTGGAGGCTCTAGAGTTTCTGTTCAACAGCT 121
 Qy 266 TyrGlyValGluValGluValGluAsnProTyrAspProAsnLeuMetValPheMet 285
 Db 122 TATGAGTGGAGTGTGAGTGTGATTAACAATCATTTGACCCGAGCTGATGTTTCATG 181
 Qy 286 AspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPhe 305
 Db 182 GATATAGAGACTATGTCTAGACACAGCTCAATCTTAGAAGCTAATATCCACATTT 241
 Qy 306 LeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAlaSer 325
 Db 242 CTCATATGCCATGCCCATGTCTCAACACAGAGTCTTTTCGAGAGAACTTGTGTTGCTCA 301
 Qy 326 LysAspAlaMetProPheAspLeuLysArgLysLeuMetSerArgLeuLysThrLeu 345
 Db 302 AAAGATGCAATGCCCATGTCTGTAAAGAAAAAAGTATGATTTGAACACCCCTC 361
 Qy 346 GlyIleGlnValThrLysValTyrGluGluTyrPheTyrIleProValGlyGlySer 365
 Db 362 GGTGTAAAGATTAAGAAATTTACGAGGAGGAATGCTTACATACCATCAGTTGGAGATCT 421
 Qy 366 LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetValHisPro 385
 Db 422 TTGCCAAATACAGACAAAAACACTTGCATTTGCTGTCTGTCGTCGATGTTGTTATCCA 481
 Qy 386 AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIle 405
 Db 482 GCCACAGGTTATTCAGTCTGATGCTGAGTCTGAGCTTGAAGCTCCAAATGCGCCCTCGTGT 541
 Qy 406 AlaIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaVal 425
 Db 542 GCAATATATACGACAAAATCATAGCAAGACATGCTTACT--AGTCAAGTACCCCG 598
 Qy 426 AsnIleSerMetGlnAlaTyrPheSerSerLeuTyrProLysGluArgLysArgGlnArgAla 445

```

Db 599 AGTATTCAACTCAAGCTGGGAACACCTTTGGCCACAGAACGAAACACATAGATCG 658
:::||||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
RESULT 4
BQ994803 661 bp mRNA linear EST 21-AUG-2002
LOCUS QGF8D16.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
DEFINITION QGF8D16, mRNA sequence.
ACCESSION BQ994803
VERSION BQ994803.1 GI:22414338
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
REFERENCE 1 (bases 1 to 661)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
JOURNAL http://compgenomics.ucdavis.edu/
COMMENT Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmall.ucdavis.edu]
belongs to contig QG_CA_Contig1411, see http://cgpdb.ucdavis.edu/
for details.
Plate: QGF8 row: D column: 16.
FEATURES
source
Location/Qualifiers
1..861
/organism="Lactuca sativa"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGF8D16"
/clone_lib="QG_EFGHJ lettuce serriola"
/lab_host="E.coli"
/note="Vector: pBRCDNASFIAB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB-QG_EFGHJ lettuce serriola
TAG_TISSUE-flowers pre-fertilized
TAG_SEQ-CCTTGACGGG"
BASE COUNT 196 a 133 c 135 g 197 t
ORIGIN
Alignment Scores:
Pred. No.: 7 49e-100 Length: 661
Score: 860.00 Matches: 167
Percent Similarity: 85.71% Conservative: 25
Best Local Similarity: 74.55% Mismatches: 28
Query Match: 31.35% Indels: 5
DB: 14 Gaps: 2
US-09-701-395a-23 (1-529) x BQ994803 (1-661)
QY 294 LysLeuGlnCysSerGluGluThrProThrPheLeuTyrrValMetProMetSerPro 313
|||::: |||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3 AAACCGGAATCTTTAGAACGAAATATCCGACTTCTCTATGTCATGCAATGCTCCA 62
QY 314 ThrArgLeuPheGluThrCysLeuAlaSerLysAspAlaMetPropheAspLeu 333

```

```

Db 63 ACAAATAATCTTCGAGGAACCTTGTTAGCTTCAGAGAACCCATGCTTCAATCTT 122
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 334 LeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTy 353
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 CTAAAGTCAAACTCATGTCACGATTAAGGCAATGGGTATCCGATAACAGAAC-- 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 354 GluGluGluTrpSerTyrrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsn 373
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 -----GAATGGTCGTATATATCCCGTAGTGGATCGTTACCTAATACAGAACAAAGAAT 232
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 374 LeuAlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrrSerValValArg 393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 CTGGCAATTTGGTCTCGAGCTAGCTAGCTGGTCCACCCCTCCACAGGATTCAGTTGCCGA 292
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 394 SerLeuSerGluAlaProLysTyrrAlaSerValIleAlaLysIleLeuLysGlnAspAsn 413
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 TCCTTTGTCAGAACTCTCTAATATGTCAGCAGCTATGCTAAGATTTTAAACACAGATCAA 352
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 414 SerAlaTyrrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSer 433
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 TCTAAAGAGATGATTTCT---CTTGAAAAATACACTAACATTTCAAACCAAGCATGGAA 409
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 434 SerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuLeu 453
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 ACATTGTGCCACTTGAAGGAAAGACACGAGCGCTTCTTCTGTTCCGACTATCACAC 469
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 454 IleValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuPro 473
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 ATCGTCTAATGATCTAGAGGGAACACGACATTTTCCGCTACTTCTTCTGTTCCGCTCC 529
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 474 ThrTrpMetTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPhe 493
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 AAATGCGATGTTGGTGGGATTTTGGGCTCTCTTTATCTTCAACGGGATTTGATAATATTT 589
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 SerMetTyrrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeu 513
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 590 GCGCTTATATGTTGTGTAGTACACCTCACAGCTTGAGATGAAGTGAACCTGGGTAGACATCTA 649
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 514 LeuSerAspPro 517
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 650 CTTTCTGATCCG 661
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 5
AV931085 719 bp mRNA linear EST 18-JAN-2002
LOCUS AV931085 K. Sato unpublished cDNA library, cv. Haruna Nijo second
DEFINITION leaf stage seedling leaves Hordeum vulgare subsp. vulgare cDNA
clone basd18h18 3', mRNA sequence.
ACCESSION AV931085
VERSION AV931085.1 GI:18226882
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 719)
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..719
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
FEATURES
source

```

/clone="basd1818"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo second leaf stage seedling leaves"
 /tissue_type="seedling leaves"
 /dev_stage="second leaf stage"
 BASE COUNT 207 a 158 c 170 g 182 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1,18e-97 Length: 719
 Score: 843.50 Matches: 164
 Percent Similarity: 84.9% Conservative: 28
 Best Local Similarity: 72.5% Mismatches: 33
 Query Match: 30.7% Indels: 2
 DB: 10 Gaps: 1

US-09-701-395A-23 (1-529) x AV931085 (1-719)

QY 303 ProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluThrCys 322
 DB 713 CCAACATTCTTATGTCATGCCATGTCATCCACAGAGTTTCTTTGAGGAACATGC 654
 QY 323 LeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeu 342
 DB 653 TTAGCTTCAAGAGTGAATGCCCTTTGATCTCTTAAAGAGGTTGATGTCGGTTG 594
 QY 343 LysThrLeuGlyIleGlnValThrLysValTyrGluGluThrPsrTyrIleProVal 362
 DB 593 GATGCGATGGGATCGTATCTATAAGTATACGAGGAGGTTGTCATTATCTCTGTT 534
 QY 363 GlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMet 382
 DB 533 GGAGGATCCTTACCTAACACACAGACAGAAAAATCTTGCTGCTGCAGCGAGTATG 474
 QY 383 ValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAla 402
 DB 473 GTCCACCTCGCACTGGATGTCGGTGGTCAGATCTTCTGAGGCTCAAGATATGCT 414
 QY 403 SerValIleAlaLysIleLeu---LysGlnAspAsnSerAlaTyrValValSerGlyGln 421
 DB 413 TCTGTGATATCTGATATCTAGGAATCGTCTATCTTGGACAATATTGCTTGGAACT 354
 QY 422 SerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLys 441
 DB 353 TCTGAATGTCAGTCCATCAATGCTGATGGGAAACACTATGGCTCAAGAACGGA 294
 QY 442 ArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAla 461
 DB 293 CGTCAACGCTATCTTCTCTTGGATTGGCTTGATAATTCAACTGGATAACGAAGC 234
 QY 462 ThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPheLeu 481
 DB 233 ATCCAAACGTTCTCGAAACCTTTTCCGGTTACCAATGATGTGGCGAGGATTCCTT 174
 QY 482 GlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAla 501
 DB 173 GGTTCGACCTCTCGTCACTGATCTCATGCTGTTGCACTTACATGTTTGCATTTGG 114
 QY 502 ProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAlaVal 521
 DB 113 CCAACAAATTTGCGAATGAATCGTGCAGACCTCTCTCGGACCCAC -GGTTCAGCA 55
 QY 522 MetValArgAlaTyrLeu 527
 DB 54 ATGATCAGGACCTACCTG 37

RESULT 6
 AV926200
 LOCUS
 DEFINITION AV926200 K. Sato unpublished cDNA library, cv. Haruna Nijo second
 leaf stage seedling leaves Hordeum vulgare subsp. vulgare cDNA
 clone basd1818 5', mRNA sequence.
 AV926200
 ACCESSION

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AV926200.1 GI:18221997

EST.

Hordeum vulgare subsp. vulgare.

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.

1 (bases 1 to 665)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasi Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 665

FEATURES

source

/organism="Hordeum vulgare subsp. vulgare"

/cultivar="Haruna Nijo"

/db_xref="taxon:112509"

/clone_lib="basd1818"

/clone_lib="K. Sato unpublished cDNA library, cv. Haruna

Nijo second leaf stage seedling leaves"

/tissue_type="seedling leaves"

/dev_stage="second leaf stage"

BASE COUNT 170 a 145 c 151 g 197 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 2,99e-91 Length: 665

Score: 793.50 Matches: 153

Percent Similarity: 81.36% Conservative: 26

Best Local Similarity: 69.55% Mismatches: 40

Query Match: 28.93% Indels: 1

DB: 10 Gaps: 1

US-09-701-395A-23 (1-529) x AV926200 (1-665)

QY 288 ArgAspTyrMetGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyr 307
 DB 4 AGATTGTTTCAAGNAGNAGTTCTCACACCTTCAGGAGCCCAATCCACATTTCTTAT 63
 QY 308 ValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAsp 327
 DB 64 GCATGGCCATGTCATCCACAGAGTTTCTTTGAGGAACATGCTTAGCTTCAAAAGAT 123
 QY 328 AlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIle 347
 DB 124 GCAATGCCCTTTCATCTCTTAAGAGAGGTTGATGCTCGTGGATGGGAGTC 183
 QY 348 GlnValThrLysValTyrGluGluGluTyrSerTyrIleProValGlyGlySerLeuPro 367
 DB 184 CGTATCTAAAAGTATACGAGGAGGAGTGGTCTTATATTCCTCTTGAGGATCCTTACCT 243
 QY 368 AsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetValHisProAlaThr 387
 DB 244 ACACAGACAGAAAAATCTTGATTTGGTGTGCGAGGAGTATGGTCCACCTGCAACT 303
 QY 388 GlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLys 407
 DB 304 GGATACCTCGGTGTCAGATCTTGTCTGAAGCTCCAGATATGCTTCTGTGATATCTGAT 363
 QY 408 IleLeu---LysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsn 426
 DB 364 ATCTTAGGAAATGCTGCTATTCTGGACAATATTGCTGGAAGTCTGAAATGTCAGT 423
 QY 427 IleSerMetGlnAlaTrpSerLeuTrpProLysGluArgLysArgGlnArgAlaPhe 446
 DB 424 CCATCAATGCTTGCATGGGAAACACTATGGCTTCAAGACGGAACGCTCAACGCTCATTC 483

```

QY 447 PheLeuPheGlyLeuGluLeuValGlnLeuAspPheGluAlaThrArgThrPhePhe 466
Db 484 TTCCTCTTTGATGGCCTTGATAATACTGAACGATGACGATCCAAACGTTCTTC 543
QY 467 ArgThrPheArgLeuProThrTyrMetTyrPheGlyPheLeuGlySerSerLeuSer 486
Db 544 GAACCTTTTCCGGTACCAATGATGATGGGAGATTCCTTGGTTCACCGCTGTCG 603
QY 487 SerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArg 506
Db 604 TCAGTGGATCTCATGCTGTTGGCATCTCTACATGTTTGCATTAATGGCCCAACAATTCGGA 663
RESULT 7
BOL16144 610 bp mRNA linear EST 22-JUL-2002
LOCUS EST601720 mixed potato tissues Solanum tuberosum cDNA clone STMDG03
5' end, mRNA sequence.
ACCESSION BOL16144
VERSION BOL16144.2 GI:21917328
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 610).
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karanycheva,S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Apr 17, 2002 this sequence version replaced gi:20168106.
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potatoc@igr.org
This clone is available through the Research Genetics, contact the
Research Genetics for further information 1-800-711-6195 or
cdna@resgen.com
Seq primer: T3.
FEATURES
source Location/Qualifiers
1. .610
/organism="Solanum tuberosum"
/cultivar="Kennebec or Binjete"
/db_xref="taxon:4113"
/clone="STMDG03"
/clone_lib="mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
tubers, or roots."
BASE COUNT 187 a 128 c 127 g 168 t
ORIGIN
Alignment Scores:
Pred. No.: 6,51e-83 Length: 610
Score: 728.50 Matches: 151
Percent Similarity: 74.89% Conservative: 25
Best Local Similarity: 64.26% Mismatches: 26
Query Match: 26.56% Indels: 35
DB: 14 Gaps: 2
US-09-701-395A-23 (1-529) x BOL16144 (1-610)
QY 206 ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu 225
Db 2 GTTCTGTATCTAACTCAAGAGTGGATAGATTGTTGAGGCCACAAATGCCACATCTT 61
QY 226 ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla 245

```

```

Db 62 GT----- 63
QY 246 AlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGlnThrAla 265
Db 64 -----AGAGTTTCTGTCTCAACAGCT 84
QY 266 TyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetValPheMet 285
Db 85 TATGGAGTGAAGTTGAGTTGATAACAATCATATGACCGACCGCTGATGGTTTTCATG 144
QY 286 AspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPhe 305
Db 145 GATTATAGAGACTATGTCAGACACGACGCTCAATCTTTAGAAAGAAACTGATTACGATTGAACACCTC 204
QY 306 LeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAlaSer 325
Db 205 CTCATGCCATGCCCATGCTCCAAACAGAGTCTTTTCGAGGAAACTTGTGTGGCTTCA 264
QY 326 LysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeu 345
Db 265 AAAGATGCAATGCCATTCGATCTGTTAAAGAAAAAACTGATTACGATTGAACACCTC 324
QY 346 GlyIleGlnValThrLysValTyrGluGluTyrSerTyrIleProValGlyGlySer 365
Db 325 GGTGTAGAATTAAGAAATTTACGAGGAGGAATGGTCTTACATACACAGTTGGAGGATCT 384
QY 366 LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro 385
Db 385 TTGCCAAATACAGAACAAAAACATTCGATTTGGTCTGCTGTACATGATGGTTCACTCA 444
QY 386 AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIle 405
Db 445 GCCACAGAGTTATTACAGTCGTCAGATCACTGCTGAGCTCCAAATCGCGCTCGTCT 504
QY 406 AlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaVal 425
Db 505 GGCAAATATTACAGCAAAATCATAGCAAGACATGCTTACT---AGTTCAAGTACCCTG 561
QY 426 AsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
Db 562 AGTATTTCACCTCAAGCTTGGAC-ACCTTTGGCCACAGAAGCA 605
RESULT 8
LOCUS BG241188 541 bp mRNA linear EST 15-FEB-2001
DEFINITION OVI_39_C08_b1_A002 Ovary 1 (OVI), Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION BG241188
VERSION BG241188.1 GI:12776261
KEYWORDS EST.
SOURCE Sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 541)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
An EST database from Sorghum: ovaries of varying immature stages
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 526
POLYA-No.

```

```

FEATURES          Location/Qualifiers
  source
    1..541
      /organism="Sorghum bicolor"
      /db_xref="taxon:4558"
      /clone_lib="Ovary 1 (OVI)"
      /note="Organ: Mix Of ovaries of varying immature stages
from 8-week-old plants; Vector: pBluescript II from Lambda
zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT      146 a  107 c  140 g  147 t    1 others
ORIGIN
Alignment Scores:
Pred. No.:      1.9e-79      Length:      541
Score:          701.00      Matches:      131
Percent Similarity: 85.47%      Conservative: 22
Best Local Similarity: 73.18%      Mismatches: 26
Query Match:    25.56%      Indels:      0
DB:             12      Gaps:      0

US-09-701-395A-23 (1-529) x BG241188 (1-541)
Qy 159 GluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaPro 178
Db 5 GAGGTTGTATCGACGCTCTGGAAGGATACATATGTTCTACCTAGACAAATAACAGGCC 64
Qy 179 ValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeu 198
Db 65 ATACTGATGGCGTCTTATGGCAGGTGCACCGTGCTGCTCCATGAGGAGTACTG 124
Qy 199 LysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGlu 218
Db 125 AGAAGATGCTATGAAGCTGGCTGACATACCTGAACCTCAAGTGCACAAAGATCAGAA 184
Qy 219 AlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeu 238
Db 185 TCTCAGATGGACACAGAGTAGTCTGTGTATAGAGGTGCTGAGATATATTGTCAGGCTT 244
Qy 239 AlaThrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyPro 258
Db 245 GCCATTGTGCTCGGGGGCAGCATCTGGTGGCTCTAGAGTATGAGTGGGGGTCT 304
Qy 259 ArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnProTyrAsp 278
Db 305 CGTGTTCGTCGACACATGTCATCGGAGTAGAAGTTGAGGTGGAAATAATCCATATGAT 364
Qy 279 ProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSer 298
Db 365 CCCAGCTTAATGGTTTTCATGGACTACAGAGATTGTTTCAAGAGGAATTCACACACT 424
Qy 299 GluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePhe 318
Db 425 GAACAAGAAATCCACACTTCTCTGTATGCTATGCCATGTCATCCACAGAGTTTTCCT 484
Qy 319 GluGluThrCysLeuAlaSerLysAspAlaMetPropheAspLeuLeuLysArgLys 337
Db 485 GAGGAACATGCTAGCTTCTAAAGATGCTATGCTCTTGTATCTACTTAAAGAAGAGG 541

RESULT 9
AW906884          AW906884          548 bp      mRNA      linear      EST 24-MAY-2000
LOCUS            EST343007 potato stolon, Cornell University Solanum tuberosum cDNA
DEFINITION
clone cSTA4H14, mRNA sequence.
ACCESSION        AW906884
VERSION          AW906884.1 GI:8071094
KEYWORDS         EST.
SOURCE           potato.
ORGANISM         Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE        1 (bases 1 to 548)

```

AUTHORS van der Hoeven, R., Bezzerides, J., Bachem, C., Horvath, B., Visser, R.,
 Holt, I. E., Liang, F., Hansen, T. S., Utterback, T., Bowman, C. L., Doan,
 B., Bougri, O., Buell, C. R., Ronning, C. M., Tanksley, S. D. and Baker,
 B.
 TITLE Generation of ESTs from potato swelling stolons
 JOURNAL Unpublished (1999)
 COMMENT Contact: Research Genetics, Libraries Division
 Tel: 1-800-711-6195
 Email: cdna@resgen.com
 5 prime sequence.
 FEATURES Location/Qualifiers
 source
 1..548
 /organism="Solanum tuberosum"
 /cultivar="Bintje"
 /db_xref="taxon:4113"
 /clone_lib="cSTA4H14"
 /clone_lib="potato stolon, Cornell University"
 /tissue_type="axillary buds of stem explants, swelling
 stolons"
 /dev_stage="1 to 3 days"
 /lab_host="SOLR"
 /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
 XhoI: RNA was supplied by Christian Bachem & Beatrix
 Horvath (Laboratory of Plant Breeding, Dept. of Plant
 Sciences, Wageningen University, The Netherlands). Total
 RNA was isolated from developing axillary buds of potato
 nodal stem cuttings cultured on medium for the
 introduction of tuber formation as described in Bachem et
 al. (Plant Journal 1996). Tissue samples were taken of
 stages corresponding to growing stolons and the early
 stages of tuber formation."
 BASE COUNT 169 a 111 c 114 g 154 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.16e-78 Length: 548
 Score: 695.00 Matches: 139
 Percent Similarity: 76.08% Conservative: 20
 Best Local Similarity: 66.51% Mismatches: 18
 Query Match: 25.34% Indels: 33
 DB: 10 Gaps: 1
 US-09-701-395A-23 (1-529) x AW906884 (1-548)
 Qy 206 ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu 225
 Db 2 GTTCTGTATCTAAACTCGAAGTGGATAGGATTGTGAGCCCAAAATGCCACAGCTCT 61
 Qy 226 ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla 245
 Db 62 GT----- 63
 Qy 246 AlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGlnThrAla 265
 Db 64 -----AGAGTTCTCTTCAACAGCT 84
 Qy 266 TyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetValPheMet 285
 Db 85 TATGAGTGGAGTGGAGTTGATTAACAATCCATATGACCCGAGCTGATGTTTCATG 144
 Qy 286 AspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPhe 305
 Db 145 GATTATAGAGATATGCTATGACACGACGCTCAATCTTTAGAACGCTAAATATCCACATTT 204
 Qy 306 LeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAlaSer 325
 Db 205 CTCTATGCCATGCCATGCTCTCAACACAGAGCTTTTTCGAGGAACACTTGTGTTGCTCA 264
 Qy 326 LysAspAlaMetPropheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeu 345
 Db 265 AAAGATGCAATGCCATTCGATCTGTAAAGAAAAAACTATGTTAGATTGAACACCTC 324
 Qy 346 GlyIleGlnValThrLysValTyrGluGluGluTyrSerTyrIleProValGlyGlySer 365

Db	325	GGTGTTAGAAATAAAGAAATTTACAGAGAGAATGGTCTTACATACCAGTTGGAGGAFTCT	384
Qy	366	LeuproAsnThrGlulnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro	385
Dd	385	TTCGCACAATCAGAACAAAAACACTTGCAATTGGTGTCTGCTAGCATGGTTCAATCCA	444
Qy	386	AlathrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValile	405
Dd	445	CCACAGGTTATTTCAGTCGTCTCAAATCACTGTCTGAAGTCCAAAATGCCCTCGTGCTT	504
Qy	406	AlalystleLeuLysGlnAspAsnSer	414
Dd	505	GCAATATATTACGACAAATCATAGC	531
RESULT	10		
BE231430			
LOCUS			
DEFINITION	BE231430	Suaeda salsa ZAP cdna library Suaeda maritima subsp. salsa	linear EST 16-OCT-2000
ACCESSION	BE231430	cDNA similar to lycopene epsilon cyclase, mRNA sequence.	
VERSION			
KEYWORDS	BE231430.1	GI:8996190	
SOURCE		Eukarya; Suaeda maritima subsp. salsa.	
ORGANISM		Suaeda maritima subsp. salsa	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;	
TITLE		Caryophyllidae; Caryophyllales; Chenopodiaceae; Suaeda.	
JOURNAL		Zhang, L., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.	
COMMENT		Expressed sequence tags from a halophyte Suaeda salsa cDNA library Unpublished (2000)	
		Contact: Hui Zhang	
		Key Laboratory of Plant Stress Research	
		The Biology Department of Shandong Normal University	
		No.88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC	
		Tel.: (86)531-2966954	
		Fax: (86)531-2966954	
		Email: zhanghesdn@edu.cn.	
FEATURES			
source		Location/Qualifiers	
		1..571	
		/organism="Suaeda maritima subsp. salsa"	
		/db_xref="taxon:126914"	
		/clone_lib="Suaeda salsa ZAP cDNA library"	
		/dev_stage="seedling"	
		/note="Organ: aerial part tissue; Vector: lambda zap;"	
		Site_1: EcoRI; Site_2: XhoI; total RNA extraction from	
		NaCl(400mM) treated Suaeda salsa by RNAsen kit(Promega);	
		mRNA isolation by MESSAGEMAKER kit(GIBCO BRL);	
		directional CDNA synthesis(EcoRI XhoI) by CDNA synthesis	
		kit(STRAVAGEN); the ZIP express library by GigapackIII	
		Gold Cloning kit(STRAVAGEN)"	
BASE COUNT	131 a	115 c	137 g
ORIGIN			188 t
Alignment Scores:			
Pred. No.:	7.16e-69	Length:	571
Score:	619.50	Matches:	117
Percent Similarity:	85.54%	Conservative:	25
Best Local Similarity:	70.48%	Mismatches:	22
Query Match:	22.58%	Indels:	2
DB:	10	Gaps:	1
US-09-701-395A-23 (1-529) x BE231430 (1-571)			
Qy	84	GlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProHillePro---	102
Dd	76	CAGATAAGGCTATGGATCTCTACTATTAATTCACAAAGCTCTCTCCATAAACAGTA	135
Qy	103	PheGlycluserValMetAspLeuValilleglYcysGlyProAlagLyLeuSerLeu	122
Dd	136	GTTGGGGACCAATTTGTTTCATATGGTGGTTGTTGGTTCGGCTCCTCGGCTTCGTTTG	195


```

Source      1. .465
/organism="Suaeda maritima subsp. salsa"
/db_xref="taxon:126914"
/clone_lib="Suaeda salsa ZAP cDNA library"
/dev_stage="seedling"
/site="Organ: aerial part tissue; Vector: lambda zap;
Site_1: EcORI; Site_2: XhoI; total RNA extraction from
NaCl(400mM) treated Suaeda salsa by RNAGENT kit(Promega);
mRNA isolation by MESSAGEMAKER kit(GIBCO BRL);
directional cDNA synthesis(EcORI XhoI) by cDNA synthesis
kit(STRATAGEN); the ZAP express library by GigapackIII
Gold Cloning kit(STRATAGENE)"
BASE COUNT 121 a 100 c 121 g 122 t 1 others
ORIGIN

Alignment Scores:
Pred. No.:      2.24e-66      Length:      465
Score:          599.00      Matches:    118
Percent Similarity: 88.51%      Conservative: 13
Best Local Similarity: 79.73%      Mismatches: 17
Query Match:    21.84%      Indels:    1
DB:             1.0      Gaps:      0

US-09-701-395A-23 (1-529) x AW982124 (1-465)
QY 187 ArgValSerArgHisLeuLeuHisGluLeuLeuLysArgCysValGluSerGlyVal 206
DB 22 AGGTCACGCCGCAATTTACTTCACCAAGAGCTCGTCAAAAGGTGTTTCGAGTCAGGGGTC 81
QY 207 SerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuVal 226
DB 82 TCTTATCTGAATGCAAAAGTCGATAGGATCATGAGGAGGTGATGATCTAGTCTTGTTN 141
QY 227 ValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaAla 246
DB 142 GAAGA-GAAGATGATATACATATTCCGTGAGGCTTCAACTGTTCATCTGGAGCAGCT 200
QY 247 SerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGlnThrAlaTyr 266
DB 201 TCAGGGAAGCTCTGGAGTAGTAAGTCTGGGGTCCAGGGTGTGTTTCAACAGCTTAT 260
QY 267 GlyValGluValGluValGluAsnProTyrAspProAsnLeuMetValPheMetAsp 286
DB 261 GGTGTGAAGTCGAGGTGGAAACAGTCTTATGATCCCAACGTGATGGTCTCATGGAC 320
QY 287 TyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeu 306
DB 321 TACAGAGATTACACCAACAGAGGTTTCAGTCTCCGAGGAGCAAGTATCCACATTTCTG 380
QY 307 TyrValMetProMetSerProThrArgLeuPhePheGluGluThrCysLeuAlaSerLys 326
DB 381 TACGCAATGCCACTGTCCACTACTAGATCTTCTTCGAGGAGACTTGTCTGGCTTCAATA 440
QY 327 AspAlaMetProPheAspLeuLeu 334
DB 441 GATGCAATGCCATTTGATCTACTC 464

RESULT 14
AV822597
LOCUS
DEFINITION AV822597 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-09-E07 5',
mRNA sequence.
ACCESSION AV822597
VERSION AV822597.1 GI:19864646
KEYWORDS EST.
SOURCE
  thale cress.
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
  1 (bases 1 to 668)
REFERENCE
  Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
  Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,

```

```

TITLE
JOURNAL
COMMENT
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4339
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
further details.
FEATURES
  Location/Qualifiers
    source
      1..668
        /organism="Arabidopsis thaliana"
        /db_xref="taxon:3702"
        /clone="RAFL05-09-E07"
        /clone_lib="RAFL5"
        /dev_stage="rosette plants"
        /lab_host="SOLR"
        /notes="Site_1: SstI; Site_2: XhoI; subjected to
        dehydration-treated(1.2, 5, 10, 24 hr)"
BASE COUNT 175 a 95 c 183 g 214 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      4.88e-66      Length:      668
Score:          598.50      Matches:    120
Percent Similarity: 74.61%      Conservative: 24
Best Local Similarity: 62.18%      Mismatches: 37
Query Match:    21.82%      Indels:    12
DB:             1.0      Gaps:      4

US-09-701-395A-23 (1-529) x AV822597 (1-668)
QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTr 16
DB 106 ATGAGTGTGTGGGCTAGGAATTTCCGAGCAATGGCAGGTTTCAACATTTCCGTCATG 165
QY 16 pThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrG1 36
DB 166 GAGTTCTCGAAGGAAATTTCCAGTGTGTTAAGAGATACAGCTATAGGAATATTCGTTTCG 225
QY 36 ySerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerAr 54
DB 226 T---TTGTGT-----AGTGTCAAGCTAGCGCGCGGAGGTTCCGGTAGTGA 270
QY 54 gSerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyG1 74
DB 271 GAGTTGTGTAGCGGTGAGAGAAGATTTCCCTGACGAAGAAGATTTGTGAAAGCTGGTG 330
QY 74 ySerGluLeuLeuPheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLe 94
DB 331 TTCTGAGATTCTATTGTTTCAAAATGCAGCAGAACAAAGATATGGATGAACATCTAAGCT 390
QY 94 uAlaAspLysLeuProPheIleProPheGlyGluSerValMetAspLeuValIleG1 114
DB 391 TGTGATAGTTGCTCTATATCAATTTGGTGTGCTGCTTGGATCTAGTGTATTGG 450
QY 114 yCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValG1 134
DB 451 TTGTGGTCTGCTGCTGTTAGCTTGGCTGCAGATCAGCTAGCTGGATTAAGATTGG 510
QY 134 yLeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLy 154
DB 511 ACTCATTTGGTCCAGATCTCTCTTTTACTTAACAATATACGTTGTTGGGAAGATGAATCAA 570
QY 154 sAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAs 174

```


THIS PAGE BLANK (USPTO)

PT New carotenoid biosynthesis genes used to provide transgenic plants
XX with altered carotenoid content
PS Claim 2; Fig 14; 134pp; English.
XX
CC The present sequence represents a lycopene epsilon-cyclase protein.
CC The enzyme is involved in the carotenoid biosynthesis pathway. The
CC specification also describes other enzymes involved in this pathway,
CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
CC hydroxylase is responsible for hydroxylating the beta-endgroup in
CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
CC of epsilon-endgroups in carotenoids. The polynucleotides of the
CC invention are used to provide transgenic plants with altered
CC carotenoid content. These plants may, for example, may be protected
CC from lethal photooxidative damage or have altered colouring.
XX
SQ Sequence 529 AA;
Query Match 100.0%; Score 2743; DB 21; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.4e-271;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELLGVNRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
Db 1 MELLGVNRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
QY 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
QY 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
Db 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
QY 241 VASGAASGLLEYEVGPRVCQTAYGVEVEVNNPDNLMVPMYRDYMQKLCQSEE 300
Db 241 VASGAASGLLEYEVGPRVCQTAYGVEVEVNNPDNLMVPMYRDYMQKLCQSEE 300
QY 301 EYPTFLVMPSPRLFFETCTCLASKDAMPDLDLKRKLSRLKTLGIQVTKVYEEENSYI 360
Db 301 EYPTFLVMPSPRLFFETCTCLASKDAMPDLDLKRKLSRLKTLGIQVTKVYEEENSYI 360
QY 361 PVGSLPNTQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDINSAYVYSG 420
Db 361 PVGSLPNTQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDINSAYVYSG 420
QY 421 QSSAVNTSMQAWSLWPKEKRRORAFFLGLLEIVQDIEATRTFFTFRLPTWMMWGF 480
Db 421 QSSAVNTSMQAWSLWPKEKRRORAFFLGLLEIVQDIEATRTFFTFRLPTWMMWGF 480
QY 481 LGSLSLSDFLVLFNMFVLPAPNSMRSLVRHLLSDPSGAVMVRAYLER 529
Db 481 LGSLSLSDFLVLFNMFVLPAPNSMRSLVRHLLSDPSGAVMVRAYLER 529
RESULT 2
AA534303
ID AA534303 standard; Protein; 529 AA.
XX
AC AA534303;
XX
DT 06-APR-2000 (first entry)
XX
DE An Adonis palaeatina lycopene epsilon-cyclase.
XX
KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;

dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
lethal photooxidative damage; colouring.
Adonis palaeatina.
Key Location/Qualifiers
FT Misc-difference 55 /note= "encoded by ACT"
FT Misc-difference 210 /note= "encoded by AAT"
FT Misc-difference 231 /note= "encoded by GAC"
FT Misc-difference 352 /note= "encoded by ATT"
FT Misc-difference 524 /note= "encoded by AAA"
XX W09963055-A1.
XX
XX 09-DEC-1999.
XX 02-JUN-1999; 99WO-US12121.
XX 02-JUN-1998; 98US-0088724.
XX 02-JUN-1998; 98US-0088725.
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX
XX Cunnigham FX, Sun Z;
XX WPI; 2000-116366/10.
XX N-PSDB; AA245595.
XX
XX New carotenoid biosynthesis genes used to provide transgenic plants
XX with altered carotenoid content
XX
XX Disclosure; Fig 24; 134pp; English.
XX
CC The present sequence represents a lycopene epsilon-cyclase protein.
CC The enzyme is involved in the carotenoid biosynthesis pathway. The
CC specification also describes other enzymes involved in this pathway,
CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
CC hydroxylase is responsible for hydroxylating the beta-endgroup in
CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
CC of epsilon-endgroups in carotenoids. The polynucleotides of the
CC invention are used to provide transgenic plants with altered
CC carotenoid content. These plants may, for example, may be protected
CC from lethal photooxidative damage or have altered colouring.
XX
SQ Sequence 529 AA;
Query Match 100.0%; Score 2743; DB 21; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.4e-271;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELLGVNRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
Db 1 MELLGVNRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
QY 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
QY 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
Db 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240

QY 241 VASGAASGKLLYEYVGGPRVCVQTAYGVEVEVENNPDNLMVFMVMDYRDYMOOKLQCSSE 300
 Db 241 VASGAASGKLLYEYVGGPRVCVQTAYGVEVEVENNPDNLMVFMVMDYRDYMOOKLQCSSE 300

QY 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKYEEWSYI 360
 Db 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKYEEWSYI 360

QY 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420
 Db 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420

QY 421 QSSAVNISMQANSSLPKPKRKQRAFFLFGLELIVOLDIEATRTFFRPLTMMWGF 480
 Db 421 QSSAVNISMQANSSLPKPKRKQRAFFLFGLELIVOLDIEATRTFFRPLTMMWGF 480

QY 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529
 Db 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529

RESULT 3
 AAY54306
 ID AAY54306 standard; Protein; 529 AA.
 XX
 AC AAY54306;
 XX
 DT 06-APR-2000 (first entry)
 XX
 DE Amino acid sequence of a lycopene epsilon-cyclase protein.
 XX
 KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring.
 XX
 OS Adonis palaestina.
 XX
 PN WO9963055-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 02-JUN-1999; 99WO-US12121.
 XX
 PR 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX, Sun Z;
 XX
 WPI; 2000-116366/10.
 XX
 DR New carotenoid biosynthesis genes used to provide transgenic plants
 XX with altered carotenoid content -
 XX
 PS Disclosure; Fig 25; 134pp; English.
 XX
 CC The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 529 AA;

Query Match 100.0%; Score 2743; DB 21; Length 529;
 Best Local Similarity 100.0%; Pred. No. 1.4e-271;
 Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVRNLISCPVWTFGRNLSSSKLAYNTHRYGSSCRVDPQVRADGGSGRSRVAYK 60
 Db 1 MELLGVRNLISCPVWTFGRNLSSSKLAYNTHRYGSSCRVDPQVRADGGSGRSRVAYK 60

QY 61 EGFVDEEDFIKAGSSELLFVQMOOTKSMKQAKLADKLPPIPGESVMDLVIGCCPAGL 120
 Db 61 EGFVDEEDFIKAGSSELLFVQMOOTKSMKQAKLADKLPPIPGESVMDLVIGCCPAGL 120

QY 121 SLAAEAAKGLKVLGILGPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
 Db 121 SLAAEAAKGLKVLGILGPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180

QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCNEIFIPORLAT 240
 Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCNEIFIPORLAT 240

QY 241 VASGAASGKLLYEYVGGPRVCVQTAYGVEVEVENNPDNLMVFMVMDYRDYMOOKLQCSSE 300
 Db 241 VASGAASGKLLYEYVGGPRVCVQTAYGVEVEVENNPDNLMVFMVMDYRDYMOOKLQCSSE 300

QY 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKYEEWSYI 360
 Db 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKYEEWSYI 360

QY 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420
 Db 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420

QY 421 QSSAVNISMQANSSLPKPKRKQRAFFLFGLELIVOLDIEATRTFFRPLTMMWGF 480
 Db 421 QSSAVNISMQANSSLPKPKRKQRAFFLFGLELIVOLDIEATRTFFRPLTMMWGF 480

QY 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529
 Db 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529

RESULT 4
 AAY54307
 ID AAY54307 standard; Protein; 529 AA.
 XX
 AC AAY54307;
 XX
 DT 06-APR-2000 (first entry)
 XX
 DE Amino acid sequence of a lycopene epsilon-cyclase protein.
 XX
 KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring.
 XX
 OS Adonis palaestina.
 XX
 PN WO9963055-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 02-JUN-1999; 99WO-US12121.
 XX
 PR 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX, Sun Z;
 XX
 WPI; 2000-116366/10.
 XX

PT New carotenoid biosynthesis genes used to provide transgenic plants
 XX with altered carotenoid content
 PS Disclosure; Fig 25; 134pp; English.
 XX
 CC The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 529 AA;
 Query Match 99.5%; Score 2728; DB 21; Length 529;
 Best Local Similarity 99.1%; Pred. No. 4.9e-270;
 Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELLGVRNLISCPWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 DB 1 MELLGVRNLISCPWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 QY 61 EGFVDEEDFKAGGSELLFVQMOTKSMKQAKLADKLPPFPFSGVMDLVVIGCGPAGL 120
 DB 61 EGFVDEEDFKAGGSELLFVQMOTKSMKQAKLADKLPPFPFSGVMDLVVIGCGPAGL 120
 QY 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAKDTIYVLNDAPVL 180
 DB 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAKDTIYVLNDAPVL 180
 QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 DB 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 QY 241 VASGAASGKLLLEYVGGPRVCQQTAYGVEVEVNNPDNLMFMDYRDYMQKQLOCSEE 300
 DB 241 VASGAASGKLLLEYVGGPRVCQQTAYGVEVEVNNPDNLMFMDYRDYMQKQLOCSEE 300
 QY 301 EYPTFLYMPSPRLFEETCLASKDAMPDCLKRLMSRLKTLGIQVTKIYEENSYI 360
 DB 301 EYPTFLYMPSPRLFEETCLASKDAMPDCLKRLMSRLKTLGIQVTKIYEENSYI 360
 QY 361 PVGSLPNTQKNLAFGAASMVHPATGYSVRSLSAPKVASIAKILKODNSAYVYVG 420
 DB 361 PVGSLPNTQKNLAFGAASMVHPATGYSVRSLSAPKVASIAKILKODNSAYVYVG 420
 QY 421 QSSAVNISMQAWSLWPKERRQRAFFLGLLELVQIDIEATRTFFRFLPTMMWGF 480
 DB 421 QSSAVNISMQAWSLWPKERRQRAFFLGLLELVQIDIEATRTFFRFLPTMMWGF 480
 QY 481 LGSLSLSDFLVFSMYMFVLAPNSMRSLVRLHLLSDPSGAVMVYAYLER 529
 DB 481 LGSLSLSDFLVFSMYMFVLAPNSMRSLVRLHLLSDPSGAVMVYAYLER 529
 RESULT 5
 AAY52593
 ID AAY52593 standard; Protein: 529 AA.
 XX
 AC AAY52593;
 XX
 DT 07-MAR-2000 (first entry)
 XX
 DE Adonis palaestina epsilon lycopene cyclase.
 XX
 KW Lycopene; epsilon cyclase; bicyclic epsilon carotene; biosynthesis;
 KW carotenoid; epsilon ring; lactucaxanthin; protein production;

KW transgenic plant; altered synthesis; overexpression;
 KW nutritional value; pharmacology; colour; cancer; antisense;
 KW accumulation; substrate; beta-carotene; vitamin A.
 XX
 OS Adonis palaestina.
 XX
 PN WO9961399-A1.
 XX
 PD 02-DEC-1999.
 XX
 PF 25-MAY-1999; 99WO-US10461.
 XX
 PR 26-MAY-1998; 98US-0084222.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX;
 XX
 DR WPI; 2000-062667/05.
 DR N-PSDB; AAZ46329.
 XX
 PT New eukaryotic epsilon-lycopene cyclase and related DNA, used to
 PT regulate carotenoid synthesis in plants and other hosts, e.g. for
 PT treatment of cancer
 PS Disclosure; Fig 6B; 38pp; English.
 XX
 CC This sequence represents epsilon lycopene cyclase from Adonis
 CC palaestina. This enzyme adds two epsilon rings to lycopene to
 CC form bicyclic epsilon-carotene. Carotenoids with two epsilon rings
 CC are found in significant amounts in relatively few plants (romaine
 CC lettuce produces an abundance of such a carotenoid, lactucaxanthin),
 CC while carotenoids with two beta rings are ubiquitous and those
 CC with one beta and one epsilon ring are common. The invention relates to
 CC epsilon lycopene cyclases capable of forming bicyclic epsilon-carotene,
 CC and also to methods of transforming known carotenoids into novel or rare
 CC products. Epsilon lycopene cyclase nucleic acids may be used for
 CC recombinant protein production and for the generation of transgenic
 CC plants that have altered patterns of carotenoid synthesis. Such plants
 CC may accumulate new or rare carotenoids or overexpress known carotenoids,
 CC resulting in altered nutritional value, pharmacology or colour.
 CC Carotenoids isolated from such transgenic plants are useful
 CC therapeutically, e.g., for treatment of cancer. Alternatively, epsilon
 CC lycopene cyclase antisense nucleic acids can be used to cause
 CC accumulation of the substrates of the enzyme, specifically beta-carotene
 CC or derived products such as vitamin A. The isolated eukaryotic enzyme
 CC allows synthesis of epsilon, epsilon-carotene (which can be
 CC isolated from natural sources only in small amounts) on a large scale,
 CC enabling bulk production of derived carotenoids.
 XX
 SQ Sequence 529 AA;
 Query Match 99.5%; Score 2728; DB 21; Length 529;
 Best Local Similarity 99.1%; Pred. No. 4.9e-270;
 Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELLGVRNLISCPWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 DB 1 MELLGVRNLISCPWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 QY 61 EGFVDEEDFKAGGSELLFVQMOTKSMKQAKLADKLPPFPFSGVMDLVVIGCGPAGL 120
 DB 61 EGFVDEEDFKAGGSELLFVQMOTKSMKQAKLADKLPPFPFSGVMDLVVIGCGPAGL 120
 QY 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAKDTIYVLNDAPVL 180
 DB 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAKDTIYVLNDAPVL 180
 QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 DB 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 QY 241 VASGAASGKLLLEYVGGPRVCQQTAYGVEVEVNNPDNLMFMDYRDYMQKQLOCSEE 300

CC carotenoids. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAT95371). Expression vector pAREPS
 CC comprising the epsilon cyclase gene is deposited as ATCC 98005.
 CC The claimed eukaryotic carotenoid biosynthetic enzymes
 CC epsilon cyclase, beta-carotene hydroxylase and isopentenyl
 CC pyrophosphate isomerase (see AAT95346-51) are used in methods for
 CC augmenting the accumulation of carotenoids and for the production
 CC of novel and rare carotenoids in host cells. Methods are also
 CC provided for controlling the ratio of various carotenoids in a
 CC host, and for screening for eukaryotic genes that encode enzymes
 CC of carotenoid biosynthesis and metabolism.
 XX
 XX Sequence 524 AA;
 SQ
 Query Match 71.4%; Score 1959.5; DB 18; Length 524;
 Best Local Similarity 70.4%; Pred. No. 2.5e-191;
 Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
 QY 1 MELLGVRNL-----ISSCPWTFGTNRNLSSKLAYNIHRYGSSCRVDFQVRADGG--SGSR 54
 DB 1 MECVGARNAFAAMAVSTFSSWCRKRPVVKRYSYRNIRFG-LC-----SVRASGGSGSGE 55
 QY 55 SSVAKEGVDEEDFIKAGGSELLFVQMOQKMSKQAKLADKLPPIPFGEVMDLVVIG 114
 DB 56 SCVAVREFDAEDFVKAGGSEILFVQMOQKMDQESKLVKLPPIGIGDGLDHHVIG 115
 QY 115 CGPAGLSAAEAAKLGKLVGLIGDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIVYLD 174
 DB 116 CGPAGLAAEAAKLGKLVGLIGDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIVYLD 175
 QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFI 234
 DB 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFI 235
 QY 235 PCRLATVASGAASGKLLLEYEVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYDMOQK 294
 DB 236 PCRLATVASGAASGKLLLEYEVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYDMOQK 295
 QY 295 LQCSSEYPTFLYVMPMSPTLRFEEETCLASKDAMPDGLKRLKMSRLKTLGIQVTKYE 354
 DB 296 VRSLEAEYPTFLYVMPMSPTLRFEEETCLASKDAMPDGLKRLKMSRLKTLGIQVTKYE 355
 QY 415 AYVVSQSSAVNISMQAWSSLWPKRQRRAFFLGLGLIQLVQDTEATRTFFRFLPK 469
 DB 416 KQINS-----NISQAWDTLWPKRQRRAFFLGLGLIQLVQDTEATRTFFRFLPK 469
 QY 475 WNWGFLGSSLSFDLVLFMYMFLAPNSMRSLVRLHLLSDPAGVMVRAYLE 528
 DB 470 WNWQFLGSLTSGDLVLFALYMEVISPNNLRKGLINHLISDPTGATIKTYLK 523
 RESULT 8
 AAY54283
 ID AAY54283 standard; Protein; 524 AA.
 XX
 XX AAY54283;
 XX
 XX 06-APR-2000 (first entry)
 XX
 DE An Arabidopsis thaliana lycopene epsilon-cyclase.
 XX
 XX Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 XX lethal photooxidative damage; colouring.
 XX
 OS Arabidopsis thaliana.
 XX
 PN W09963055-A1.

XX 09-DEC-1999.
 PD
 XX 02-JUN-1999; 99WO-US12121.
 PF
 XX 02-JUN-1998; 98US-0088724.
 PR
 XX 02-JUN-1998; 98US-0088725.
 PR
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX Cunningham FX, Sun Z;
 PI
 XX WPI; 2000-116366/10.
 DR
 XX N-ESDB; AAZ45578.
 DR
 XX New carotenoid biosynthesis genes used to provide transgenic plants
 PT with altered carotenoid content -
 PT
 XX Example; Page 64-65; 134pp; English.
 PS
 XX The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 XX Sequence 524 AA;
 SQ
 Query Match 71.4%; Score 1959.5; DB 21; Length 524;
 Best Local Similarity 70.4%; Pred. No. 2.5e-191;
 Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
 QY 1 MELLGVRNL-----ISSCPWTFGTNRNLSSKLAYNIHRYGSSCRVDFQVRADGG--SGSR 54
 DB 1 MECVGARNAFAAMAVSTFSSWCRKRPVVKRYSYRNIRFG-LC-----SVRASGGSGSGE 55
 QY 55 SSVAKEGVDEEDFIKAGGSELLFVQMOQKMSKQAKLADKLPPIPFGEVMDLVVIG 114
 DB 56 SCVAVREFDAEDFVKAGGSEILFVQMOQKMDQESKLVKLPPIGIGDGLDHHVIG 115
 QY 115 CGPAGLSAAEAAKLGKLVGLIGDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIVYLD 174
 DB 116 CGPAGLAAEAAKLGKLVGLIGDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIVYLD 175
 QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFI 234
 DB 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFI 235
 QY 235 PCRLATVASGAASGKLLLEYEVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYDMOQK 294
 DB 236 PCRLATVASGAASGKLLLEYEVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYDMOQK 295
 QY 295 LQCSSEYPTFLYVMPMSPTLRFEEETCLASKDAMPDGLKRLKMSRLKTLGIQVTKYE 354
 DB 296 VRSLEAEYPTFLYVMPMSPTLRFEEETCLASKDAMPDGLKRLKMSRLKTLGIQVTKYE 355
 QY 355 EWSYIPVGGSLPNTQKNLAFGAASMVHPATGYSVVRSLSEAPKYASVIAKILKQNS 414
 DB 356 EWSYIPVGGSLPNTQKNLAFGAASMVHPATGYSVVRSLSEAPKYASVIAKILRETT 415
 QY 415 AYVVSQSSAVNISMQAWSSLWPKRQRRAFFLGLGLIQLVQDTEATRTFFRFLPK 474
 DB 416 KQINS-----NISQAWDTLWPKRQRRAFFLGLGLIQLVQDTEATRTFFRFLPK 469
 QY 475 WNWGFLGSSLSFDLVLFMYMFLAPNSMRSLVRLHLLSDPAGVMVRAYLE 528
 DB 470 WNWQFLGSLTSGDLVLFALYMEVISPNNLRKGLINHLISDPTGATIKTYLK 523


```

XX SQ Sequence 516 AA;
Query Match 70.5%; Score 1934; DB 21; Length 516;
Best Local Similarity 71.4%; Pred. No. 1e-188;
Matches 365; Conservative 64; Mismatches 74; Indels 8; Gaps 2;

QY 22 NLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGFVDEDFIKAGSELLFVQ 81
DB 7 HMTATMAAFTCPWFMTSIRYTKQIKC---NAAKSQLVVKQETEEEDYKAGSELLFVQ 63
QY 82 MQOTSMKQAKLADKLPPIPFG-----ESVMDLVVIGCGPAGLSAAAPAKLGLKVLGI 136
DB 64 MOONKMDAQSSLSQKLPVPVIGGGSDSNCILDLVVIGCGPAGLAGESAKLGLNVALI 123
QY 137 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVLTGRAYGRVSRHLLHEE 196
DB 124 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVLTGRAYGRVSRHLLHEE 183
QY 197 LKRCVSGSVYLDKVERITEAGDGHSLVVCENEIFIPCLATVAGSAAGSKLLEYEYG 256
DB 184 LLTRCMESGVSYLSKVERITEAPNGLSLIECEGNITIPCLATVAGSAAGSKLLEYEYG 243
QY 257 GPRVCVQTAYGVEVEVENNPDPNLMVMDYRDYMQQKLCSEEEYPTFLYVMPSPTRL 316
DB 244 GPRVCVQTAYGVEVEVESTPYDPSLMVMDYRDYTKHKSQSLEAQYPTFLYVMPSPTKV 303
QY 317 FFEETCLASKDAMPFDLLKRLKMSRLKTLGIVTKVYEEESYIPVGGSLPNTKOKNLAF 376
DB 304 FFEETCLASKDAMPFDLLKRLKMSRLKTLGIVTKVYEEESYIPVGGSLPNTKOKNLAF 363
QY 377 GAAASVMHPATGYSVVRSLSEAPKYASVIAKILKQDNISAYVYVSGSSAVNISMQAWSSLW 436
DB 364 GAAASVMHPATGYSVVRSLSEAPKYASVIAKILKQDNISAYVYVSGSSAVNISMQAWSSLW 423
QY 437 PKERRQRQAFLLGLELIVOLDIEATRTFFRFRPLPTWMMWFGSSLSFDFLVFSMY 496
DB 424 PLERRQRQAFLLGLELIVQMDIEGTRFTFFRFRPLPTWMMWFGSSLSFDFLIIFAFY 483
QY 497 MFVLAPNSMRSLVRHLLSDPSCGAVMVRAYL 527
DB 484 MFIIAPHSLRWGLVRHLLSDPTGGTMLKAYL 514

RESULT 11
ID AAY54310 standard; Protein; 516 AA.
XX AC AAY54310;
XX DT 06-APR-2000 (first entry)
XX DE Amino acid sequence of a lycopene epsilon-cyclase protein.
KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
KW lethal photooxidative damage; colouring.
XX OS Tagetes erecta.
XX PN WO9963055-A1.
XX PD 09-DEC-1999.
XX PF 02-JUN-1999; 99WO-US12121.
XX PR 02-JUN-1998; 98US-0088724.
XX PR 02-JUN-1998; 98US-0088725.
XX PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX PI Cunningham FX, Sun Z;

```

```

XX DR WPI; 2000-116366/10.
XX PT New carotenoid biosynthesis genes used to provide transgenic plants
XX with altered carotenoid content -
XX PS Disclosure; Fig 25; 134pp; English.
XX CC The present sequence represents a lycopene epsilon-cyclase protein.
XX CC The enzyme is involved in the carotenoid biosynthesis pathway. The
XX CC specification also describes other enzymes involved in this pathway,
XX CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
XX CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
XX CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
XX CC hydroxylase is responsible for hydroxylating the beta-endgroup in
XX CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
XX CC of epsilon-endgroups in carotenoids. The polynucleotides of the
XX CC invention are used to provide transgenic plants with altered
XX CC carotenoid content. These plants may, for example, may be protected
XX CC from lethal photooxidative damage or have altered colouring.
XX SQ Sequence 516 AA;
Query Match 70.5%; Score 1934; DB 21; Length 516;
Best Local Similarity 71.4%; Pred. No. 1e-188;
Matches 365; Conservative 64; Mismatches 74; Indels 8; Gaps 2;

QY 22 NLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGFVDEDFIKAGSELLFVQ 81
DB 7 HMTATMAAFTCPWFMTSIRYTKQIKC---NAAKSQLVVKQETEEEDYKAGSELLFVQ 63
QY 82 MQOTSMKQAKLADKLPPIPFG-----ESVMDLVVIGCGPAGLSAAAPAKLGLKVLGI 136
DB 64 MOONKMDAQSSLSQKLPVPVIGGGSDSNCILDLVVIGCGPAGLAGESAKLGLNVALI 123
QY 137 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVLTGRAYGRVSRHLLHEE 196
DB 124 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVLTGRAYGRVSRHLLHEE 183
QY 197 LKRCVSGSVYLDKVERITEAGDGHSLVVCENEIFIPCLATVAGSAAGSKLLEYEYG 256
DB 184 LLTRCMESGVSYLSKVERITEAPNGLSLIECEGNITIPCLATVAGSAAGSKLLEYEYG 243
QY 257 GPRVCVQTAYGVEVEVENNPDPNLMVMDYRDYMQQKLCSEEEYPTFLYVMPSPTRL 316
DB 244 GPRVCVQTAYGVEVEVESTPYDPSLMVMDYRDYTKHKSQSLEAQYPTFLYVMPSPTKV 303
QY 317 FFEETCLASKDAMPFDLLKRLKMSRLKTLGIVTKVYEEESYIPVGGSLPNTKOKNLAF 376
DB 304 FFEETCLASKDAMPFDLLKRLKMSRLKTLGIVTKVYEEESYIPVGGSLPNTKOKNLAF 363
QY 377 GAAASVMHPATGYSVVRSLSEAPKYASVIAKILKQDNISAYVYVSGSSAVNISMQAWSSLW 436
DB 364 GAAASVMHPATGYSVVRSLSEAPKYASVIAKILKQDNISAYVYVSGSSAVNISMQAWSSLW 423
QY 437 PKERRQRQAFLLGLELIVOLDIEATRTFFRFRPLPTWMMWFGSSLSFDFLVFSMY 496
DB 424 PLERRQRQAFLLGLELIVQMDIEGTRFTFFRFRPLPTWMMWFGSSLSFDFLIIFAFY 483
QY 497 MFVLAPNSMRSLVRHLLSDPSCGAVMVRAYL 527
DB 484 MFIIAPHSLRWGLVRHLLSDPTGGTMLKAYL 514

RESULT 12
ID AAY54308 standard; Protein; 533 AA.
XX AC AAY54308;
XX DT 06-APR-2000 (first entry)
XX DE Amino acid sequence of a lycopene epsilon-cyclase protein.

```


XX	WO9963055-A1.
PN	
XX	
PD	09-DEC-1999.
XX	
PD	
XX	02-JUN-1999; 99WO-US12121.
PF	
XX	
XX	02-JUN-1998; 98US-0088724.
PR	
XX	02-JUN-1998; 98US-0088725.
XX	
XX	
PA	{UYMA-} UNIV MARYLAND BALTIMORE.
XX	
XX	Cunningham FX, Sun Z;
PI	
XX	
XX	WPI; 2000-116366/10.
DR	
XX	
PT	New carotenoid biosynthesis genes used to provide transgenic plants
PT	with altered carotenoid content -
XX	
PS	Disclosure; Fig 25; 134pp; English.
XX	
CC	The present sequence represents a lycopene epsilon-cyclase protein.
CC	The enzyme is involved in the carotenoid biosynthesis pathway. The
CC	specification also describes other enzymes involved in this pathway,
CC	such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
CC	beta-carotene hydroxylase. IPP isomerase catalyses the reversible
CC	conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
CC	hydroxylase is responsible for hydroxylating the beta-endgroup in
CC	carotenoids. Lycopene epsilon-cyclase is responsible for the formation
CC	of epsilon-endgroups in carotenoids. The polynucleotides of the
CC	invention are used to provide transgenic plants with altered
CC	carotenoid content. These plants may, for example, may be protected
CC	from lethal photooxidative damage or have altered colouring.

XX New carotenoid biosynthesis genes used to provide transgenic plants
PT with altered carotenoid content -
XX
PS Disclosure; Fig 25; 134pp; English.

The present sequence represents a lycopene epsilon-cyclase protein. The enzyme is involved in the carotenoid biosynthesis pathway. The specification also describes other enzymes involved in this pathway, such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and beta-carotene hydroxylase. IPP isomerase catalyses the reversible conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene hydroxylase is responsible for hydroxylating the beta-endgroup in carotenoids. Lycopene epsilon-cyclase is responsible for the formation of epsilon-endgroups in carotenoids. The polynucleotides of the invention are used to provide transgenic plants with altered carotenoid content. These plants may, for example, may be protected from lethal photooxidative damage or have altered colouring.

Query Match 69.6%; Score 1909.5; DB 21; Length 526;
Best Local Similarity 68.5%; Pred. No. 3.3e-186;
Matches 368; Conservative 68; Mismatches 78; Indels 23;

QY	1	MELLGVNRNLSSCPV-----WTFGRNLSSK---LAYNTHRYGSSCRVDFQVRADGG	50
Db	1	MECVGVQNVGAMAVLTPRLNRWSSG--ELCQEKSIFFLAY--EQYESKC-----NSS	48
QY	51	SGSRSSVAYKEGVDEEDFIKAGGSELLFVQMQOOTSMEKQAKLADKLPIIPFGESVMDL	110
Db	49	SGSDSCVVKEDFADBEDYIKAGGSOLVFVQMQKKDMDQQSKLSDDELQOISAGQTVL	108
QY	111	VWTCGCPAGLSLAAEAAKGLKVLGICGPDLPFTNNTGVWDEDFKDLGLERCITEHAWKDTI	170
Db	109	VWTCGCPAGLALAAESAKGLNVLGICGPDLPFTNNTGVWDEDFKDLGLCQACIEHVWRTI	168
QY	171	VYLDNDAPVLICRAYGRVSRHLLHEELLKRCVSEGSYILDSKVERITEAGDGHSLVCVEN	230
Db	169	VYLDNDPPTIIGRAYGRVSRHFTHEELLKRCVFAGVYILNSKVDRTVEATNGSILVECGS	228

OS *Lycopodium* sp.

Db 468 RVPKMWQGLGSSUSSADLMLFAFYMFIIAPNDMRKGLIRHLLSDPTGATLIRTYL 524

RESULT 15

AAE09798 AAE09798 standard; Protein; 517 AA.

XX AC AAE09798;

XX DT 10-DEC-2001 (first entry)

XX DE Spinach lycopene epsilon cyclase protein.

XX KW Spinach; lycopene epsilon cyclase; lutein; therapeutic; alpha-carotene;

XX KW nutritional supplement; food product.

XX OS Spinacia oleracea.

XX PN WO200166703-A1.

XX PD 13-SEP-2001.

XX PF 07-MAR-2001; 2001WO-US071178.

XX PR 07-MAR-2000; 2000US-0187576.

XX PA (CRGI) CARGILL INC.

XX PI De Souza ML, Kollman SR, Schroeder WA;

XX DR WPI; 2001-582274/65.

XX DR N-PSDB; RAD17465.

XX PT Novel isolated lycopene epsilon cyclase polypeptide, useful for

XX PT producing lutein, feed supplement and enhanced food products -

XX PS Claim 8; Fig 3; 37pp; English.

XX CC The invention relates to spinach lycopene epsilon cyclase protein and
 their corresponding cDNA molecule. Lycopene epsilon cyclase is useful for
 producing lutein in microorganisms. Lycopene epsilon cyclase is also used
 as nutritional supplement or as sources of lutein for producing enhanced
 food products or other therapeutic purposes. The invention also relates
 to a method for producing lutein, where a single polypeptide converts
 alpha-carotene to lutein in the microorganism. The present sequence is
 spinach lycopene epsilon cyclase protein.

XX SQ Sequence 517 AA;

Query Match 69.0%; Score 1892; DB 22; Length 517;
 Best Local Similarity 71.4%; Pred. No. 2e-184;
 Matches 355; Conservative 61; Mismatches 75; Indels 6; Gaps 1;

Qy 37 SSCRVDFQVRADGGSGSRSS-----VAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEK 90

Db 19 SNVVECCSSRRSGSVLRANSDDSCVTAPEDFANEEDEFIKAGGSELLYVQMQQNKANDC 78

Qy 91 QAKLADKLPPFPFSGVMDLVVIGCGPAGLSLAEEAKGLKVLIGDPLDPTNNYGVWE 150

Db 79 YKSIDKLRIQISDANELLDMVYVIGCGPAGLALAAESAKLGLKVLIGDPLDPTNNYGVWE 138

Qy 151 DEFKDLGLERCIEHAKDTIYVLNDAPVLIGRAYGRVSRHLLHEELKRCVSGSVYLD 210

Db 139 DEFALGCGGCIERVWRDTIYIIDDNDPIYIGRSTGKVRQLLHKELVHRCLESGSVYLN 198

Qy 211 SKVERITEAGDGHSLVVCNEIFIPCRLATVASGAAGKLLLEYEYVGGPRVCVQTAAYGEV 270

Db 199 AKVENIMEGPDGHRVLVACERGVTPICRLVTASGAAGKLLLEYEYVGGPRVCVQTAAYGEV 258

Qy 271 EVENNPDPNLMVMDYRDYVQKQKQCSEEEYPTFLYVMPSPTRLFFETCLASKDAMP 330

Db 259 EVENSPDPNVMVMDYRDYTKLSVQSLKAYPTFLYAMPISPTRIFFETCLASVDAMP 318

Qy 331 FDLKKRLMSRLKLTGIGVTKVYEEWSYIPVGGSLPNTQKNLAFGAASVWHPATGYS 390
 Db 319 FDLAKKLMTRLOTMGVTRITKIYEENSYIPVGGSLPNTQORNLAFAAASVWHPATGYS 378
 Qy 391 VVRSLEAPKYASVIAKILKODNSAYVYVSGSSAYVNIQMAWSSLPKRRORAFFLFG 450
 Db 379 VVRSLEAPKYASAIANLIKNDLSKNAILQRQSVGNISQMAWNTLWPOERKQRAFFLFG 438
 Qy 451 LELIVOLDIEATRTFFRTFFRLPTMMWGFGLGSSLSLSSFDLVLFMYMFVLAAPNSMRMSLV 510
 Db 439 LSLIVOLDIEGIRTEFFRTFFRVKMWEGFLGSLSSADLILFAFYMFIIAPNDLRMGILI 498
 Qy 511 RHLSDPSGAVMVRAYL 527
 Db 499 RHLSDPTGATMIRIYI 515

Search completed: May 21, 2003, 22:09:01
 Job time : 68 secs

THIS PAGE BLANK (USPTO)


```

QY 55 SSVAYKEGVDEEDPIKAGGSELLFVQMOQKSMKQAKLADKLPPIPFPGSVMDLVVIG 114
Db 56 SCVAVREDFADEEDFVKAGGSEILFVQMOQKMDQKSLVDKLPPIGSDGALDHVYIG 115
QY 115 CGPAGLSAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKOTIVYLD 174
Db 116 CGPAGLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFDGLQKCIHVVWRETIVYLD 175
QY 175 NDAPVLIGRAYGRVSRHLLHEELKRCVSGSVYLDKSRVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRVSRHLLHEELKRCVSGSVYLDKSRVERITEAGDGLRLVACDDNNVI 235
QY 235 PCRLATVAGSAAKGLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOK 294
Db 236 PCRLATVAGSAAKGLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOK 294
QY 295 LOCSEERYPTLYVMPMSPTLRFEEETCLASKDAMPDGLLKRKLSRLKTLGIQVTKYE 354
Db 296 VRSLEAEPTFLYAMPMTKSLRFEETCLASKDAMPDGLLKRKLSRLKTLGIQVTKYE 355
QY 355 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIAKILKODNS 414
Db 356 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIAKILREET 415
QY 415 AYVYSGSSAVNISQAWSSLPKRRQRAFFLGLGLIYVOLDIEATRTFFRFLPT 474
Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLGLGLIYVOLDIEATRTFFRFLPK 469
QY 475 WMMGFLGSSLSFDLVLFMYMFLAPNSMRSLVRLHLLSDPSGAVMVRAYLE 528
Db 470 WMMGFLGSLTSLGDLVLFALYMFVISPNNLRKGLNHLISDPTGATWIKYILK 523

```

RESULT 2

```

us-08-624-125-21
; Sequence 21, Application US/08624125
; Patent No. 574341
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
; TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/624,125
; FILING DATE: 29-MAR-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: KELBER, STEVEN B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2747-063-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
us-08-624-125-21
Query Match 71.4%; Score 1959.5; DB 1; Length 524;
Best Local Similarity 70.4%; Pred. No. 1e-200;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
QY 1 MELLGVRNL----ISSCPVTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGG--SGSR 54
Db 1 MECVGARNAFAAMAVSTFPSSWSCRKRPVVKYSYRIRFG--LC-----SVRASGSGSSGSE 55
QY 55 SSVAYKEGVDEEDPIKAGGSELLFVQMOQKSMKQAKLADKLPPIPFPGSVMDLVVIG 114
Db 56 SCVAVREDFADEEDFVKAGGSEILFVQMOQKMDQKSLVDKLPPIGSDGALDHVYIG 115
QY 115 CGPAGLSAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKOTIVYLD 174
Db 116 CGPAGLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFDGLQKCIHVVWRETIVYLD 175
QY 175 NDAPVLIGRAYGRVSRHLLHEELKRCVSGSVYLDKSRVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRVSRHLLHEELKRCVSGSVYLDKSRVERITEAGDGLRLVACDDNNVI 235
QY 235 PCRLATVAGSAAKGLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOK 294
Db 236 PCRLATVAGSAAKGLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOK 294
QY 295 LOCSEERYPTLYVMPMSPTLRFEEETCLASKDAMPDGLLKRKLSRLKTLGIQVTKYE 354
Db 296 VRSLEAEPTFLYAMPMTKSLRFEETCLASKDAMPDGLLKRKLSRLKTLGIQVTKYE 355
QY 355 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIAKILKODNS 414
Db 356 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIAKILREET 415
QY 415 AYVYSGSSAVNISQAWSSLPKRRQRAFFLGLGLIYVOLDIEATRTFFRFLPT 474
Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLGLGLIYVOLDIEATRTFFRFLPK 469
QY 475 WMMGFLGSSLSFDLVLFMYMFLAPNSMRSLVRLHLLSDPSGAVMVRAYLE 528
Db 470 WMMGFLGSLTSLGDLVLFALYMFVISPNNLRKGLNHLISDPTGATWIKYILK 523
RESULT 3
us-09-201-641-6
; Sequence 6, Application US/09201641A
; Patent No. 6232530
; GENERAL INFORMATION:
; APPLICANT: Cunningham Jr, Francis X
; APPLICANT: Dellapenna, Dean
; TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
; FILE REFERENCE: Quest 41-162
; CURRENT APPLICATION NUMBER: US/09/201,641A
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Tagetes erecta
us-09-201-641-6
Query Match 70.5%; Score 1934; DB 4; Length 516;
Best Local Similarity 71.4%; Pred. No. 5.3e-158;
Matches 365; Conservative 64; Mismatches 74; Indels 8; Gaps 2;
QY 22 NLSSSKLAYNIHRYGSSCRVDFQVRADGGSSSVAYKEGVDEEDFVKAGGSELLFVQ 81
Db 7 HMTATMAAFTCPRPMTSIRVTQKIC---NAAKSQLVVKQIEEEDYVKKAGGSELLFVQ 63
QY 82 MQQTKSMEKQAKLADKLPPIPFPG-----ESVMDLVVIGCGPAGLSAAEAAKGLKVLGI 136

```


Db 78 GRAQDFVILIIICAGPAGRLAEOVSKYGIKVCVDPSP-PLSMWPNNYGVWDFENLGLD 136

Qy 161 CIEHAKDITVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVYLSKVERITEAG 220

Db 137 CLDHKPMTCVHINDNKTIVLGRYGRVSRHLLKLLNSCVENRVRFKYAKWKV-EHE 195

Qy 221 DGHSLVVCNEIFPCRLATVASCAGSKLLEYEGGPR-VCVQTAIXGVEVEENNYPDP 279

Db 196 EFESSIVCDDGKKIRGLSVNDASGFAS-DFIEYD--KPRNHGYQIAHGVLEVDNHPFDL 252

Qy 280 NLAYFMYRDY---MOQKLOCSEEEYPTFLYVMPSPTRLFFETCLASKDAMPFDLLKR 336

Db 253 DKVLMWRDROSHLGNPEYLRVNNAKEPTFLYAMPDFDNLVLEETSLVSRPVLSEYMEVKR 312

Qy 337 KIMSRKLTGLQVTKVVEEESYIPVGGSLPNTQKNLAFGAASVMHPATGYSVVRSL 396

Db 313 RNVARLHLGIVKRSVIEEKCVPIMGGLPRIPQNVMAIGNSGIVHPSTGYMVARSMA 372

Qy 397 EAPKVASVIKILKODNSAYVVGSSAVNISQAWSLPKKRRQRAFFLFGLELIVQ 456

Db 373 LAPVLAEAIVEGL--GSTRMIRGS---QIYHRVWNGLWPLDRRCVRECYSGMEYLLK 425

Qy 457 LDIEATRTFFTRFLPMTWMMGFLGSSLSFDLVLFMSYMFVLAPNSMRMSLV 510

Db 426 LDLKGTRLDFAFDLPDKYQWGLSSRLSVKELGLLSLCLFHHGHSNLTRLDIV 479

RESULT 6

US-09-134-607A-17

Sequence 17, Application US/09134607A

Patent No. 6252141

GENERAL INFORMATION:

APPLICANT: Joseph Hirschberg et al.

TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION

TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE

TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID

TITLE OF INVENTION: BIOSYNTHESIS

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

STREET: 2001 Jefferson Davis Highway, Suite 207

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead, Slimnote 890TX

OPERATING SYSTEM: MS DOS version 6.2,

OPERATING SYSTEM: Windows version 3.11

SOFTWARE: Word for Windows version 2.0,

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/134.607A

FILING DATE:

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Friedman, Mark M.

REGISTRATION NUMBER: 33,883

REFERENCE/DOCKET NUMBER: 325/12

TELEPHONE: 972-3-5625553

TELEFAX: 972-3-5625554

TELEX:

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 498

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-134-607A-17

Query Match 26.2%; Score 719.5; DB 4; Length 498;

Best Local Similarity 40.1%; Pred. No. 4.7e-68;

Matches 164; Conservative 78; Mismatches 148; Indels 19; Gaps 9;

Qy 109 DLVVIGCGPAGLSAABAALGLKVLIGLIDPLPT---NNYGVWEDEKDLGLERCIEHA 165

Db 83 DVIIGAGPAGRLAEOVSKYGIKVCVDPSP-PLSMWPNNYGVWDFENLGLDCLDK 141

Qy 166 WKOTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVYLSKVERITEAGDGHSL 225

Db 142 WPMTCVHINDNKTIVLGRYGRVSRHLLKLLNSCVENRVRFKYAKWKV-EHEEFP 200

Qy 226 VVCNEIFPCRLATVASCAGSKLLEYEGGPR-VCVQTAIXGVEVEENNYPDPNLV 284

Db 201 IVCDGKKIRGLSVNDASGFAS-DFIEYD--KPRNHGYQIAHGVLEVDNHPFDL 257

Qy 285 MDYRDY---MOQKLOCSEEEYPTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLMSR 341

Db 258 MDWRDHLGNPEYLRVNNAKEPTFLYAMPDFDNLVLEETSLVSRPVLSEYMEVKRMV 317

Qy 342 LKTLGIVTKVVEEESYIPVGGSLPNTQKNLAFGAASVMHPATGYSVVRSLSEAPKY 401

Db 318 LRHLGIVKRSVIEEKCVPIMGGLPRIPQNVMAIGNSGIVHPSTGYMVARSMA 377

Qy 402 ASVTAKILKODNSAYVVGSSAVNISQAWSLPKKRRQRAFFLFGLELIVOLDIEA 461

Db 378 AEAIVEGL--GSTRMIRGS---QIYHRVWNGLWPLDRRCVRECYSGMETLLKLDLKG 430

Qy 462 TRTFRTRFFTRFLPMTWMMGFLGSSLSFDLVLFMSYMFVLAPNSMRMSLV 510

Db 431 TRLDFAFDLPDKYQWGLSSRLSVKELGLLSLCLFHHGHSNLTRLDIV 479

RESULT 7

US-09-134-607A-18

Sequence 18, Application US/09134607A

Patent No. 6252141

GENERAL INFORMATION:

APPLICANT: Joseph Hirschberg et al.

TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION

TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE

TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID

TITLE OF INVENTION: BIOSYNTHESIS

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

STREET: 2001 Jefferson Davis Highway, Suite 207

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead, Slimnote 890TX

OPERATING SYSTEM: MS DOS version 6.2,

OPERATING SYSTEM: Windows version 3.11

SOFTWARE: Word for Windows version 2.0,

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/134.607A

FILING DATE:

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Friedman, Mark M.

REGISTRATION NUMBER: 33,883

REFERENCE/DOCKET NUMBER: 325/12

TELEPHONE: 972-3-5625553

TELEFAX: 972-3-5625554

TELEX:


```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,009
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-147-009-11

Query Match 5.1%; Score 139; DB 4; Length 503;
Best Local Similarity 23.2%; Pred. No. 6.3e-06;
Matches 124; Conservative 66; Mismatches 184; Indels 160; Gaps 27;

QY 89 EKQAKLADKLPPIPGSGVMDLVVIGCGPAGLSLAABAAKGLKVLGIPDL--PFTNNY 146
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 30 KKVAKLPDAATEVR-RDGDADVIIVGAGVGSALAYALAKDGRVHVIERDMREPVR-- 85
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 147 GWDEDF-----KDLGLERCIE-----HAWKDTIVYL-----DNDAPV-LI 181
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 86 --MMGEFMQCGRLLLSKLGLEDCEIGDIQIATGLAYKDGQKALVSFPEDNDFPEPT 143
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 182 GRAY--GRVSRHL-----LHEELKRCVE-----SGVSLDSKVERITEAGDGH 223
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 144 GRAFYNGRFVQRLRQKASSLPTVQLEGTVKSLEEKGVIKGYTKNS-----AGE-- 194
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 224 SLVVCENEIPIPCRLAIVASGAAS-----GKILEYEVGGPRVCVOTAYGVEVE 273
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 195 -----ETTAFAP--LTVCDCGYSNLRSSVNDNNAEIVISYQVG-----YVSKNCLE-- 239
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 274 NNPDNPLMFMVDYRDYMQKLCSEEEYPTFLYVPMSPSTRLEFTECTASKDAM---- 329
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 240 ----DPEKLKLMKSPFTMLYQISSYDVRVCMEIFPGNIPISNGEMAYLKNMTAQQV 295
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 330 PFDLLKRLMSRLKTLGIQV-----TKVYEEWSPYIPVGGSLPNTPEQKNLAFGAASMVHP 385
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 296 PPEL--RKIFLKGIDECAQIKAMPTKMEATLS-----EKQGVIVLGDFAFNHRP 343
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 386 --ATGYSV-----RSLSEAPKYASVIKILKQDMSAYV--SGOSSAYNIS 428
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 344 AIASGMNVLSDILILRLLQLPLNLSLANKVSEVI-----KSFYVIRKPSMATVNTL 396
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 429 MQAWSSLW-----PKERKQRAFFLFGLELIVQLDIEATPTFTFRL-----P 473
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 397 GNAPSQVLIASDRAKAMQGCCDYL-----SSGGPRTSGMALLGGMNRP 444
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 474 TMMWGFGLSSLSFDLVLFSTMYMFLAPNSMRSLVRLHLLSDPSGAVMVRAYL 527
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 445 LSLIFHLCGILTSSIGQLSPFPSPGLGIWHSRLFGVSQMLSPAYAAAYRKSYM 498
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

RESULT 12
US-08-095-726-14
; Sequence 14, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago

```

```

;
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,566
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 530188val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-095-726-14

Query Match 4.9%; Score 134.5; DB 1; Length 374;
Best Local Similarity 20.8%; Pred. No. 1.2e-05;
Matches 87; Conservative 61; Mismatches 167; Indels 103; Gaps 16;

QY 109 DLVVIGCGPA-GL---SLAAEAAKGLKVLGIPDLPTNNYGVWEDEF---KDLGLERC 161
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 3 DLIIUGGLANGLLAWRLRQRYPOLNLLIEAGEQPGGNHWTSHEDDLTPGQAWLAPL 62
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 162 IEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELKRCVESGVSLDSKVERITEA-- 219
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 63 VAHAWPGYEQFPD-----LRRRLARG--YYSITSERFAEALH 98
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 220 ---CDGSLVVCENEIPIPCRLATV-----ASGAA--SKLLEYEVGGPRVCVOTAYG 267
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 99 QALGE-----NIWLNCVSEVLPNSVRLANGEALLAGAVIDGRVTTASSAMQTGYQ 149
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 268 VEVEVE---NNPYDPLMFMVDYRDYMQKLCSEEEYPTFLYVPMSPSTRLEFTECTLA 324
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 150 LFLGQQWRLTQPHGLTVPILMDATVAQQGYR-----FVYTLPLSADTLIEDTRYA 201
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 325 SKDAMPEDLLKRLKMSRLKTLGIQVTKVYEEWSPYIP-----VGGSLPNTPEQKNLAFGAAA 380
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 202 NVPQRDDNALRQTVDYAHSGKWLAQLEREETGCLPTTWRTVSRLCGPMRRRAASGMRA 261
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 381 SMVHPATGYS-----VVRSLSEAPKYASVIKILKQDMSAYVYVGGSSAVNISMQAWSS 434
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 262 GLFHTTGTGYSPLAVALADAIADSPRLGSV-----PLYQLTRQFAER 303
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

QY 435 LWPKE---RRQRQRAFFLFGLELIVQLDIEATPTFTFTRFFRLPTWMMWGFGLSSLSFD 489
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 304 HWRKQGFRLNRLMFLAGRE-----ENRWRYMQRYFGLPEPTVERFYAGRLSLFD 354
;| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

RESULT 13
US-08-096-623A-14
; Sequence 14, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Biosynthesis of zeaxanthin and

```

TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts

NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Welsh & Katz, Ltd.
STREET: 120 S. Riverside Plaza, 22nd Floor
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,623A
FILING DATE: 22-JUL-1993

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/805,061
FILING DATE: 09-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,921
FILING DATE: 28-FEB-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/562,674
FILING DATE: 03-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/525,551
FILING DATE: 18-MAY-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/487,613
FILING DATE: 02-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: AMO-006.1

TELEPHONE: (312) 655-1500
TELEFAX: (312) 655-1501
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 374 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-096-623A-14

Query Match
Best Local Similarity 4.9%; Score 134.5; DB 1; Length 374;
Matches 87; Conservative 61; Mismatches 167; Indels 103; Gaps 16;

QY 109 DLVWIGCGPA-GL---SLAAEAAKGLKVLGIPDLPNTNNGVWDEF--KDLGLERC 161
DB 3 DLIVGGGLANGLIANRLRQYQPNLLIIEAGOPGNGHTWSPFHEDDLTPCGHAWLAPL 62

QY 162 IEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEA-- 219
DB 63 VAHWPGEVQFPD-----LRRRLARG--YYSITSERFAEALH 98

QY 220 ---GDGHSILVCEINEIFIPCLATV-----ASGAA--SGKLLYEYVGGPRVCVQTAYG 267
DB 99 QALGE-----NWLNCSEVSELPNSVRLANGEALLAGAVIDGRGVTAASSAMQTGYQ 149

QY 268 VEVEVE---NNPYDNLMAVMDRYDMQKLCQSEEEYPTLYVMPSPTRLFFETCL 324
DB 150 LFLGQWRLTPHGLTVPILMDAIVAAQQGYR-----FVYLPLSADTLLIEDTYA 201

QY 325 SKDAMPFOLLKRLKMSRLTKTIGIQVTKYEEWSYIP-----VGGSLPNTQKNLAFGAA 380
DB 202 NVPQRDDNALRQTVTDYAHSGKWLAQLREBTGCLPTITWRVTSRLCGPMRRRAASGMRA 261

QY 381 SMVHPATGYS-----VVRSLSEAPKYASVIKILKQDNSAYVVGSSAVNISMQAWSS 434
DB 262 GLFHPTTGYSLSPLAYALADAIADSPRLGSV-----PLYQLTRFAER 303
QY 435 LWPKE---RKQRAPFLFGLLEIVQLDIEATRTFTFFRFLPTMMWGLGSSLSFSD 489
DB 304 HWRQGFRLNMLFLAGRE-----ENRWRVMQRFYGLPEPTVERFYAGRLSLFD 354

RESULT 14
US-08-663-310-6
Sequence 6, Application US/08663310
Patent No. 5811273
GENERAL INFORMATION:
APPLICANT: MISA, No. 5811273ihiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIWARA, Susumu
APPLICANT: YOKOYAMA, Akihiko
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF.
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 386 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-663-310-6

Query Match
Best Local Similarity 4.7%; Score 129; DB 2; Length 386;
Matches 92; Conservative 56; Mismatches 165; Indels 108; Gaps 18;

QY 109 DLVWIGCGPAG--LSLAAEAAKGLKVLGLI---GPDLPFTNNYGVWDEFKDLG----- 157
DB 4 DVLVLAGLANGLIATLALRAAPDLRLVLLDHAAGPSDGH-----WSCHDPDLSPDLA 58

QY 158 -LERCIEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEELLKRCVE-----SGVSYLDSK 212
DB 158 -LERCIEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEELLKRCVE-----SGVSYLDSK 212

Db 59 RLKPLRRANWPDQEVFRPHARRL-ATGYGSLDGAALADAVVRSGAEIRWDSIALLDQA 117
QY 213 VERIT-----EAGDGHSLVVCNEIFIPCLRLATVASGAASGLLEVEYGGPRVCVQTAYG 267
Db 118 GATLSCGTTRIAG-----AVLDGRGAQPSRHITVG-----FQKFGV 153
QY 268 VEVEVENNPDNLMVMDYRDYMOQKLCSEEEYPTFLYVMPMSPTLFFETCLASKD 327
Db 154 VEIETDRPHGVPRPMI-MD-----AVTQODGYRFIYLLFPFSTRILIEDTRYSDGG 204
QY 328 AMPFDLLKRLKMSRLKTLGIQVTKVYEEWSPV-----GGSLPNTPEQKN 373
Db 205 DLDDDLAALAAASHDYARQQGWTGAEVRRER-GILPIALAHDAAGFWADHAAGPVP----- 257
QY 374 LAFGAAASMVHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVVSQGS-----SAVNIS 428
Db 258 --VGLRAGFFHPTGYSL-----PYAAQVAD-----VWAGLSGPPGTDLALRGA 298
QY 429 MOAWSSLMWPKERKRORAFLEGLLEIVOLDIEATRTFFTRFLPTMMWGMFLGSSLSF 488
Db 299 IRDYAI---DRARRDRFLRLNMLFRGCAPDRRYTLLQRFYRMPHGLIERFYAGRLSVA 355
QY 489 D 489
Db 356 D 356

RESULT 15
US-09-006-491-6
; Sequence 6, Application US/09006491
; Patent No. 5972690
; GENERAL INFORMATION:
; APPLICANT: MISA, No. 59726901hiko
; APPLICANT: KONDO, Keiji
; APPLICANT: KAJIWARA, Susumu
; APPLICANT: YOKOYAMA, Akihiro
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,310
; FILING DATE: 23-SEP-1996
; APPLICATION NUMBER: WO PCT/JP94/02220
; FILING DATE: 26-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-235917
; FILING DATE: 05-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-348737
; FILING DATE: 27-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-006-491-6

Query Match 4.7%; Score 129; DB 2; Length 386;
Best Local Similarity 21.9%; Pred. No. 4.7e-05;
Matches 92; Conservative 56; Mismatches 165; Indels 108; Gaps 18;

QY 109 DLVITGCGPAG--LSLAAEAALGLKVGLI---GPDLPFTNNYGVWEDEFKDLG----- 157
Db 4 DVLLAGAGLANGLIALALRAAPDLRLVLLDHAAGPSDGH-----WSCHDDPLSDWLA 58
QY 158 -LERCIEHAWKDTIVYLDNDAPVLIIGRAYGRVSRHLLHEELKRCVE---SGVSYLDSK 212
Db 59 RLKPLRRANWPDQEVFRPHARRL-ATGYGSLDGAALADAVVRSGAEIRWDSIALLDQA 117
QY 213 VERIT-----EAGDGHSLVVCNEIFIPCLRLATVASGAASGLLEVEYGGPRVCVQTAYG 267
Db 118 GATLSCGTTRIAG-----AVLDGRGAQPSRHITVG-----FQKFGV 153
QY 268 VEVEVENNPDNLMVMDYRDYMOQKLCSEEEYPTFLYVMPMSPTLFFETCLASKD 327
Db 154 VEIETDRPHGVPRPMI-MD-----AVTQODGYRFIYLLFPFSTRILIEDTRYSDGG 204
QY 328 AMPFDLLKRLKMSRLKTLGIQVTKVYEEWSPV-----GGSLPNTPEQKN 373
Db 205 DLDDDLAALAAASHDYARQQGWTGAEVRRER-GILPIALAHDAAGFWADHAAGPVP----- 257
QY 374 LAFGAAASMVHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVVSQGS-----SAVNIS 428
Db 258 --VGLRAGFFHPTGYSL-----PYAAQVAD-----VWAGLSGPPGTDLALRGA 298
QY 429 MOAWSSLMWPKERKRORAFLEGLLEIVOLDIEATRTFFTRFLPTMMWGMFLGSSLSF 488
Db 299 IRDYAI---DRARRDRFLRLNMLFRGCAPDRRYTLLQRFYRMPHGLIERFYAGRLSVA 355
QY 489 D 489
Db 356 D 356

Search completed: May 21, 2003, 22:12:39
Job time : 28 secs

THIS PAGE BLANK (USPTO)